Journal of
Experimental Psychology

VOL. 31, NO. 1 JULY, 1942

THE EFFECT OF INTERRUPTION, COMPLETION, AND FAILURE UPON THE ATTRACTIVENESS OF ACTIVITIES

BY DORWIN CARTWRIGHT

State University of Iowa

Considerable systematic advance has been made in the psychology of motivation through use of the technique of interruption. Several investigations (**1, 6, 10, 15, 16, 19**) have demonstrated that an individual tends to resume an interrupted activity at the earliest opportunity, and other experiments (**29, 20, 26, 27, 17**) have found that interrupted activities are better recalled than completed ones. Considered together these studies provide strong support for the theory proposed by Lewin (**13, 29**) which coordinates a tension system within the person to an intention to perform a certain activity.[1] It is postulated that a psychological force acting on the person to move toward the goal either overtly or in thought accompanies the tension system and in this way there arises the tendency to resume and to recall interrupted activities. Such a tension system is discharged when the goal of the activity is reached.

Intrinsic to the concept of tension is the notion of spread of tension. According to the theory, when the tension in one system becomes markedly greater than that in another, it will spread to the second system (depending upon the permeability of the boundary between the two systems). With the second system under tension, a second force will act upon the person in the direction of the activity coordinated to the second system, establishing a tendency for the person to perform this activity also. Certain descriptive observations (**4, 8, 9**) give support to this expectation.[2]

[1] I wish to express here my gratitude to Professor Kurt Lewin for many valuable hours of discussion concerning the material presented in this paper.

[2] It is important to keep in mind that the present study is concerned with attractiveness (valence) and leaves open the question of satisfaction (value). Dembo (**4**) and Sliosberg (**28**) have found that activities which have strong valence may prove to have little value.

In working out the detailed application of the theory of need tension certain refinements and amplifications have become possible. (1) The mere occurrence of interruption is not a sufficient condition for the emergence of a tension system. Such a system will exist after interruption only if the progress toward a goal, defined in terms of the subject's psychological situation, is interrupted before that goal is reached (**17**). (2) Studies on the level of aspiration (**5**) have shown that the completion of an activity may result in the establishment of a new goal in regard to the same activity. Thus, after completion a subject may set up as his goal the performance of the same activity in a quicker time or more efficient manner than before. When such a factor is operative the subject may appear to resume the activity which he has just completed whereas he may consider it a new activity. (3) The studies of interruption cited above were performed under conditions where interruption and completion were 'neutral' to success and failure. Rosenzweig (**23, 21**) and Nowlis (**18**) have shown that the introduction of success and failure into the interruption experiments produces quite different and rather complicated results.

The present experiment was designed to gain further insight into the effects of interruption and completion, especially as they are influenced by success and failure. Two general questions are raised: What is the effect of interruption, completion, and failure upon the attractiveness of the activities thus treated? What are the effects of these experimental treatments upon the attractiveness of activities which are similar to or different from the ones immediately affected?

Procedure

When the subject entered the experimental room he found twelve tasks lying on a long table and after a short preliminary conversation he was asked to work on some of these tasks.[3] By experimental manipulation three different terminations were given to the subject's performances: one task was interrupted before the subject could complete it; another task was completed and the subject was informed that he had done all right; and a third activity was also completed but the subject was now told that he had done very poorly. The three terminations are designated as *interruption*, *completion*, and *failure* respectively.

To measure the effect of these treatments upon the attractiveness of the various tasks, the subject was required to make a preference-rating of all the tasks before and after working on each one. A comparison of the rating of a given task before and after applying an experimental treatment provided a measure of the effect of that treatment on the attractiveness of that particular task. By this method, it was possible not only to measure changes in the attractiveness of the performed tasks but also to measure these changes for every task lying on the table.

Before the study was started a model set of instructions was prepared and the experimenter thoroughly familiarized himself with it. The general course of the experiment can be grasped in finer detail from this model set of instructions.

Experimenter: "Have you heard anything about this experiment from anyone? Well, I'll

[3] At this point the subject was not told how many of the tasks he was to perform. Each subject performed three tasks in the course of the experimental session.

tell you briefly what the purpose is so that you will know why I ask you to do the things that I ask you to do. The purpose of this experiment is to study the relation between the quality of performance on an activity and the interest one has in doing that activity. In other words, do people tend to do better on tasks they want to perform? There are several tasks over there on the table. I want you to do them as quickly and as well as you can. Part of the time the activity to be performed will be selected on a basis of your choice and part of the time it will be determined by mine. I shall time you to see how long it takes you to finish the task."

The subject is shown the tasks [4] and told the purpose of each one. He is not permitted to try to work on any at this time.

Experimenter: "So that I can know what activities you prefer to do I have made out a scale and I want you to rate each task on the scale. You will notice that the scale ranges from plus 100 through 0 to minus 100. If you dislike doing a task very much you will place it beside minus 100. If, on the other hand, you like to do one very much, you will place it beside plus 100. A task which doesn't affect you one way or the other will be placed beside 0. You may place the tasks in any arrangement that seems to suit your momentary preferences. If none of them appeals to you at all it is perfectly permissible to put them all below 0. Or if you would not dislike doing any of them they may all go above 0. In making your rating I am not at all interested in long-range or enduring interests. For my purposes I only want to know which task you want to do next. Your preference may vary from one time to another. If you change your mind after you have worked on some of the tasks you will have an opportunity to indicate it on another list later. Just indicate how you feel about the tasks at this time."

The subject makes his first rating.

Experimenter: "Well, for the first one suppose you try this one." [5]

The first task is interrupted, completed, or failed according to a prearranged schedule.[6] Let us assume that it was completed.

Experimenter: "You did 'all right' on that one.[7] Your time was about average. Now that you have done one of the tasks you may have changed your mind about the rating. I don't care at all whether you are consistent or whether you change your mind. I simply want to be sure that I know what your preferences are just before you do the next task. Will you please make out another rating of the tasks as they appeal to you at this moment? Please include on your rating the task you have just finished."

The subject makes his second rating on a second rating-sheet.

Experimenter: "I'd like for you to work on this now."

The second task is submitted to another of the three treatments according to the prearranged schedule. Let us assume that it was interrupted.

Experimenter: "I'm sorry to bother you, but I find that I have made a mistake. I did not intend for you to work on that one now. Let's put it back on the table just as it is. If you happen to work on it later you may start where you left off, and I shall record the time you have spent already." [8]

The task is returned to the table on which the other tasks are located.

Experimenter: "Will you please make out another rating for me? I must be sure that I have an accurate rating of your wishes. Please include all of the tasks."

The subject makes his third rating.

Experimenter: "Would you mind trying this one now?"

---

[4] For a description of the tasks see Appendix.

[5] See Appendix for method used in selecting the tasks.

[6] For a description of the manner in which this schedule was constructed see Appendix.

[7] This condition has been designated 'completion' simply for convenience. The subject was told that he had done 'all right' in an effort to keep the situation as neutral as possible. Actually the whole experimental setting emphasized success and failure so strongly that it was impossible to create a completely neutral setting in this way.

[8] One may question the legitimacy of suggesting to the subject the possibility that he resume. Two reasons forced the inclusion of this statement. First, it was necessary to make clear to the subject that the interrupted task was to be included on subsequent ratings of attractiveness. Secondly, in preliminary experiments without the inclusion of this sentence many subjects thought that the experimenter's mistake had 'spoiled' that task so that it was no longer an available one. The sentence was carefully worded so that the possibility of resumption was clear, but every effort was made not to suggest that resumption was in any way expected.

The third task is submitted to the remaining treatment, *e.g.* failure.

Experimenter: "You finally finished that one, did you? I am afraid you didn't do very well on that. You must have got stuck somewhere or else you found that task very hard because you took —— minutes (a number close enough to the subject's actual time to avoid suspicion) while the longest time before you was only —— minutes (a number considerably smaller but still reasonable)."

"Now I want you to make out another rating for me. Please be sure you let me know what your preferences are right at this moment. Include the task you just worked on along with the others."

After completing this preference-rating the subjects were told that time did not permit working on the remaining tasks and each subject was requested to rate all of the tasks on a basis of their *similarity* to each of the performed tasks. After these ratings were completed the experimenter asked a set of questions selected to yield further information concerning the subject's reactions to the experiment.

Participating in the experiment were 35 undergraduate students at Radcliffe and Harvard Colleges. A schedule sheet soliciting their cooperation was circulated in the class of Introductory Psychology and those who were willing to volunteer became subjects.

Two ways of computing the changes in attractiveness could be followed in this study. The first consists of a computation of the number of subjects whose ratings changed in a given way. All of the graphs in this paper are based on this type of computation. The second method is to compute the *average amount of change* for the subjects. For many purposes such a measure would be preferable to the first since it gives an indication of the magnitude of the effect within each subject. It became evident in the preliminary experiments, however, that the top and the bottom of the scale served as special points of reference for the subjects and that they viewed the changes of the rating in terms of the distance to the top or bottom (*e.g.* 'half-way to the top'). Since the different tasks had different positions on the first rating, such changes, when psychologically equivalent, would yield quite different absolute scores. A more meaningful measure would, therefore, take into account the amount of change possible in the direction in which the change occurred. A ratio of the amount of change observed to the amount of change possible could be computed for each score and this value could be expressed as a percentage; but since this statistical population would be one of ratios, computations of significance, based on the *t*-test, would be subject to some doubt. For this reason, the second method of measurement is used only as an indication of the magnitude of the effects studied within subjects. It is interesting to note, in passing, that no strikingly different conclusions could be found in this experiment through the two methods of computation.

Results

1. *Effect of Interruption.*—In Fig. 1 are shown the changes in the attractiveness of the interrupted tasks. It is seen that 42 percent of the subjects raised the rating of the interrupted task, that 22 percent lowered it, and that 36 percent rated the task the same after interruption as before. A statement of the average *amount* of change will provide an indication of the intensity of the effect. Those interrupted tasks which became more attractive rose, on the average, 49 percent of the distance to the top of the preference-scale; those which became less attractive dropped, on the average, 20 percent of the distance to the bottom of the scale.

These findings seem, at first glance, to be inconsistent with the previous studies of interruption. Ovsiankina, E. Katz, and others have reported the frequency of resumption after interruption to be approximately 85 percent. Here we find an increase in attractiveness in only 42 percent of the cases and an actual *decrease* in 22 percent. How did this discrepancy arise?

There are at least two important differences between the previous studies and this one. The first concerns the rather high percentage of tasks (36 percent) in this study which remained unchanged on the preference-rating. In this experiment interruption was effected by telling the subjects that the experimenter had not intended to give them that particular task at that time; the interruption appeared to be caused by a mistake on the part of the experimenter. Many subjects tried to rectify this mistake by acting as if it had not been made, attempting to reproduce the previous rating as accurately as possible. Had a method of interruption been employed which did not tend to rigidify the rating, it is almost certain that fewer tasks

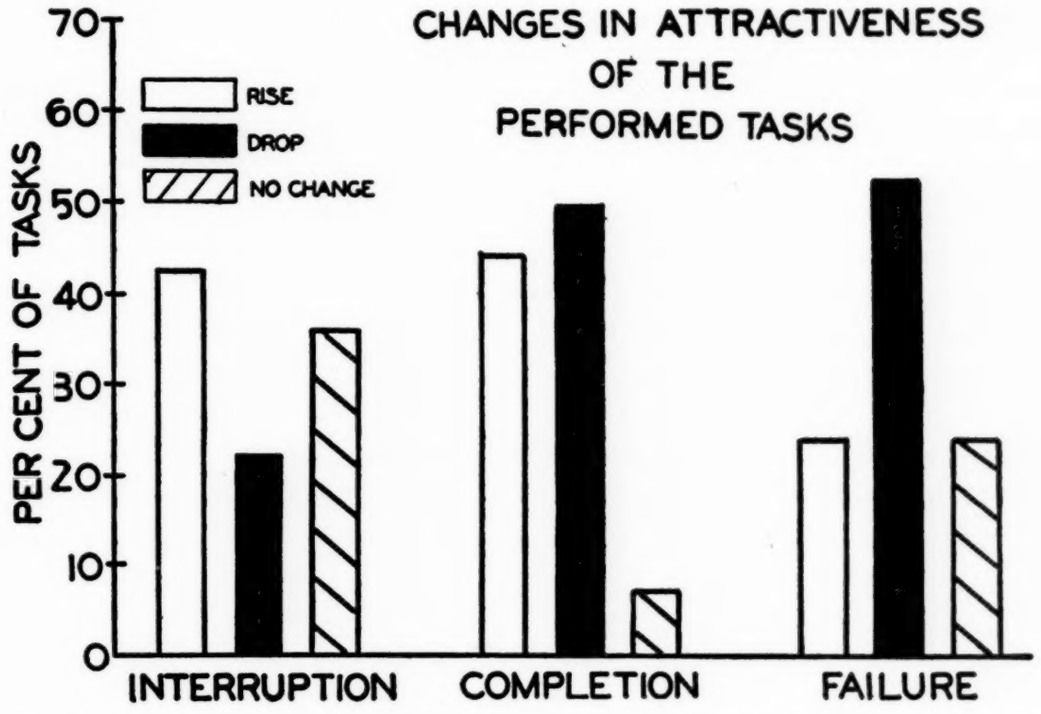

FIG. 1. Percent of performed tasks which were raised, lowered, or left unchanged on the preference-scale following each of the experimental treatments.

would have remained unchanged and that the number of tasks rated higher would have increased.

The second difference relates to those tasks whose attractiveness decreased after interruption. Spontaneous remarks by the subjects lead the experimenter to believe that some subjects viewed the interruption as an escape from certain failure. These subjects seemed to want to avoid the interrupted task. To investigate this possibility an interview was introduced at the end of the experimental session and eighteen of the subjects were asked whether at the point of interruption they had anticipated success or failure. The changes in attractiveness occurring in these two groups were distributed among the subjects as follows:

	N	Increase	Decrease	No change
Anticipation of success	11	7	1	3
Anticipation of failure	7	0	4	3

Although there are too few cases to apply a valid statistical test of significance, there can be little doubt that anticipation of failure was largely responsible for the decreases in attractiveness which followed interruption. Only one subject who anticipated success lowered the rating and no subject who anticipated failure raised it.

It cannot be stated too strongly that future studies dealing with the effects of interruption, whether concerned with resumption or with memory, cannot fail to pay strict attention to the method of interruption and to the subject's expectations of success or failure.

The theory of tension outlined at the beginning of this paper holds that whenever the tension in one system becomes sufficiently great it will spread to systems which are in dynamic communication with it. Expressed in terms of the present experiment, this statement means that whenever the attractiveness of an interrupted activity increases sufficiently there should be a concomitant increase in attractiveness of any other activities which are in communication with the first. While several factors undoubtedly determine the communication between systems, there is ample experimental evidence from studies on recognition and recall (**11**, **12**), on substitution (**15**), and on perception (**11**) that perceived similarity is one very important factor producing communication. From these considerations one should expect, then, that for an increase in the attractiveness of one activity, there should be a concomitant increase in the attractiveness of activities which are perceived as closely similar to it.

From the ratings of similarity described above it was possible to classify the tasks into two groups: (1) tasks which the subject rated as of the same type as (highly similar to) the performed task, and (2) tasks which he rated as of types different from the performed task. Fig. 2 presents the changes in attractiveness found in these two groups of tasks. When the interrupted task became more attractive, 57 percent of the tasks similar to it also became more attractive while only 15 percent became less attractive. A chi-square test shows this difference to be significant at less than the 1 percent level. Tasks of types which were rated as different from the interrupted task, on the other hand, rose and dropped about equally often (29 and 26 percent).

Most of the subjects for whom the interrupted task became less attractive, it will be recalled, indicated that they anticipated failure when they were interrupted. This fact is consistent with the findings of the studies on the level of aspiration which indicate that the attractiveness of a given activity depends upon the probability that the activity can be successfully completed. Our findings here indicate

further that the effects of failure adhere not to a single activity but rather to a category of activities. Subjects remarked, for example, "I don't seem to be very good at jig-saw puzzles" or "Mathematical problems always seem to get me." The inclusiveness of these categories undoubtedly varies from subject to subject, and it is important for an understanding of the general effect of failure upon a person to know how inclusive the category is in which the failure is experienced.[9] In relation to the present experiment tasks which were seen by the subject to be of the same category should be affected similarly by failure. Fig. 2 shows that when the interrupted task dropped,

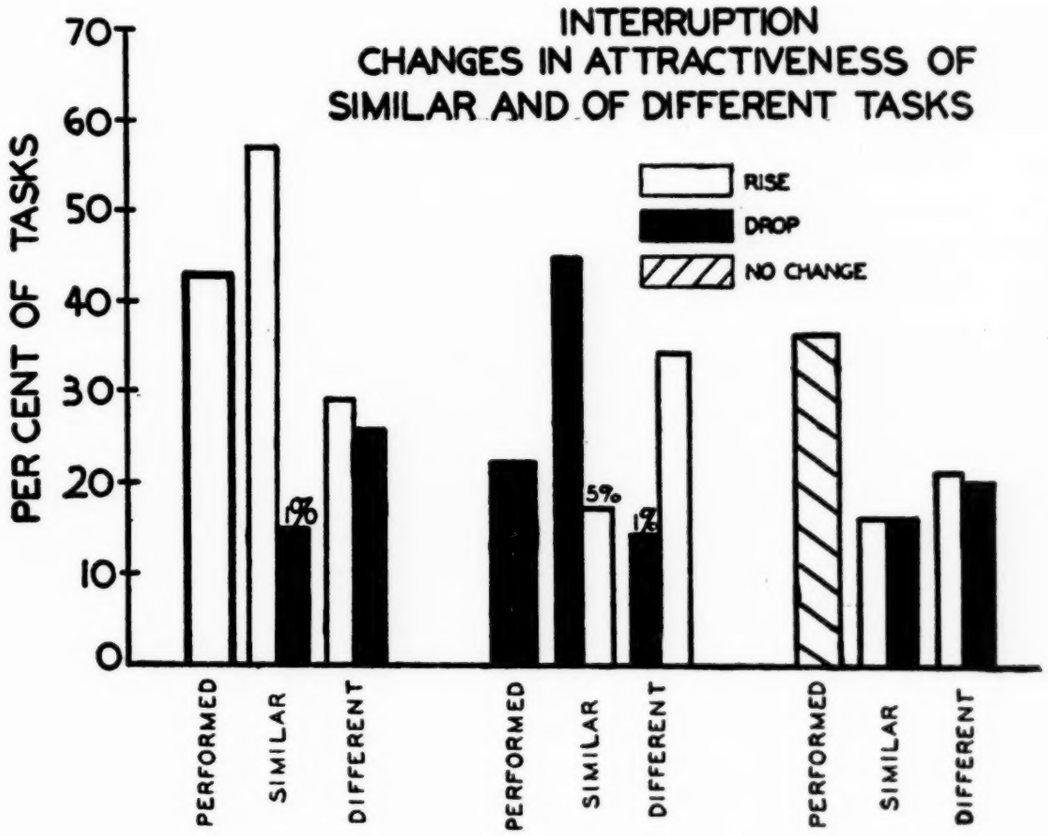

FIG. 2. Percent of similar and different tasks which were raised or lowered on the preference scale following interruption of the performed task. The tasks are grouped according to the direction of change of the interrupted task. Beginning at the left the figure should be read as follows: "When the interrupted task rose (as it did in 42 percent of the cases) 57 percent of the tasks similar to it also rose and 15 percent dropped, while 29 percent of the tasks different from it rose and 26 percent dropped." The small percent values written above the bars indicate the probability that so large a difference would occur by chance.

activities similar to it dropped more than twice as frequently as they rose. A difference as large as this would occur by chance less than five times in a hundred.

[9] A method of dealing with these categories conceptually and one possible operational definition of their boundaries are developed in a previous article (3).

Consideration of the changes in attractiveness of activities different from the interrupted task which became less attractive demonstrates a further type of interdependence of activities. When the interrupted activity dropped, activities different from it *rose* twice as frequently as they dropped, the difference being significant at the 1 percent level. A change in the attractiveness of activities in one direction is here accompanied, in other activities, by a change in the opposite direction. Fear of failure appears not only to have reduced the attractiveness of the 'dangerous' activities, but also to have positively increased the attractiveness of activities which are viewed as different. The theoretical implications of such a process certainly deserve further study.[10]

Those cases in which the interrupted task did not change on the preference-rating may be taken as a sort of control. Tasks similar to it rose and dropped infrequently and equally often; tasks of different types also showed no predominant direction of change. These facts indicate that consistent changes in the attractiveness of the related tasks are dependent upon changes in the performed tasks.

2. *Effect of Completion.*—When interpreting the effects of completion upon the attractiveness of activities, it must be kept in mind that the subject's performance was said by the experimenter to be of about average quality. Whether the subject viewed his performance as successful or unsuccessful, therefore, depended upon his own level of aspiration (7). It was clear, however, from the subjects' overt behavior and remarks that many of them felt rather satisfied with their performance; others, while not feeling failure, were a bit disppointed that they had not performed better.

In Fig. 1 it is seen that 93 percent of the completed tasks changed in attractiveness in some way, 44 percent of the tasks becoming more attractive while 49 percent became less attractive. The tasks which became more attractive rose, on the average, 33 percent of the distance to the top of the scale. Those which became less attractive dropped, on the average, 48 percent of the distance to the bottom of the scale.

If, in accordance with the theory of tension, completion results in a release of tension, why did 44 percent of the tasks become more attractive after completion? To answer this question one must turn to the studies of the level of aspiration. These studies show that, when tasks are graded according to difficulty, subjects usually select a more difficult task after successfully completing one. In this ex-

[10] This type of interdependence does not seem to be confined to the effects of failure. It is probably true, although the techniques employed in this experiment do not show it, that a need leads to both positive and negative valences. When a person is hungry, for example, not only does food become attractive, but physical work also becomes unattractive. Such a view of needs, if established, will have far-reaching effects upon motivational theory.

periment there was no clear differentiation of the tasks according to difficulty, but since the performance on each task was timed, it was possible for the subject to establish as a new level of aspiration the completion of the same task in a shorter time. It is believed, then, that in those cases where the completed task became more attractive the subjects were attempting to better their previous performance. If this proposal is correct, a new goal, and thus a new tension system, were set up following completion as a result of a shift in the level of aspiration. The changes in attractiveness, accordingly, should be essentially the same as those following interruption when success was anticipated. An inspection of Fig. 3 shows this expectation to be confirmed. When the completed tasks rose, 67 percent of the similar tasks also rose while only 11 percent of the similar tasks dropped.

The general interpretation of completion in terms of level of aspiration can also be applied to those subjects for whom the completed task became less attractive. Almost all of these subjects commented spontaneously that it would be 'meaningless' to repeat the completed task. When questioned further the subjects replied that they felt so sure of success that repeating the task would be mere routine. The decrease in attractiveness of the completed task, thus, is in line with the findings concerning the level of aspiration that completing 'too easy' tasks does not give a feeling of success and that such tasks are therefore not attractive.

This certainty of success in the future, with its resulting decrease in the attractiveness of the completed task, results from the subject's view of his own past. Past experience with particular activities affects his appraisal of the probability of success or failure upon future performance and therefore influences the attractiveness of these activities. Here, just as in the case of failure, this influence of past experience on the psychological future affects a whole category of activities rather than a single task: an experience of success increases the probability of future success in a whole area of activities. The tasks of the same cateogry as the completed one, were not similar enough to the performed task, however, to insure *certainty* of success, but, at the same time, their probability of success was increased. Accordingly, when the completed task was lowered on the preference-scale tasks of the same type were raised (Fig. 3).

A second influence of past success may also be noted in this situation. It may be seen in Fig. 3 that the affective tone of the whole experiment was raised by the success: even tasks different from the completed one rose more often than they dropped. Altogether approximately 40 percent of *all* of the tasks in the experiment became

more attractive following this successful completion, while only slightly more than 20 percent became less attractive.

3. *Effect of Failure.*—The term 'failure' as employed in this section refers to the setting where the subject, after completing the activity, was told that he had taken a longer time to finish than any other subject. Fig. 1 shows the changes in the attractiveness of these tasks on which failure was experienced. Slightly more than half of the tasks became less attractive, almost a quarter of them became more attractive, and the remaining ones were left unchanged.

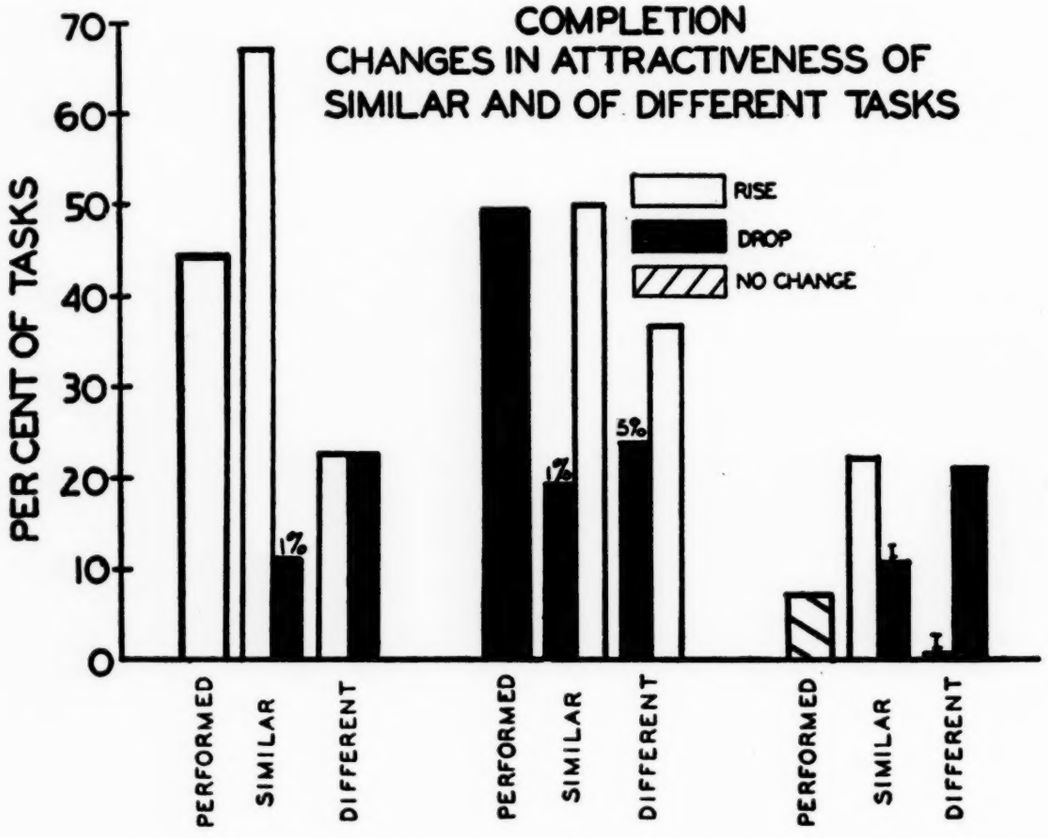

FIG. 3. Percent of similar and different tasks which were raised or lowered on the preference-scale following completion of the performed task. This figure should be read in the same manner as Fig. 2. The letter *I* indicates an insufficient number of cases to apply a chi-square test.

Those tasks which became less attractive dropped, on the average, 36 percent of the distance to the bottom of the scale. The tasks which became more attractive rose, on the average, 47 percent of the distance to the top of the scale.

It is seen in Fig. 4 that the tasks similar to the performed tasks changed predominantly in the same direction as the performed tasks and that the tasks different from them showed no preponderant direction of change. When the performed task rose, tasks similar to it rose more than twice as frequently as they dropped, the

difference being significant at the 5 percent level. When the performed task dropped, the tasks similar to it also dropped nearly four times as frequently as they rose, this difference being significant well below the 1 percent level. When the performed task was unchanged, both the tasks similar to it and different from it also remained relatively unchanged.

These data indicate that, for the largest group of subjects, failure acted to reduce the attractiveness of the tasks on which the failure was experienced. Further, just as when failure was *anticipated*, the reduction of attractiveness affected a whole category of activities.

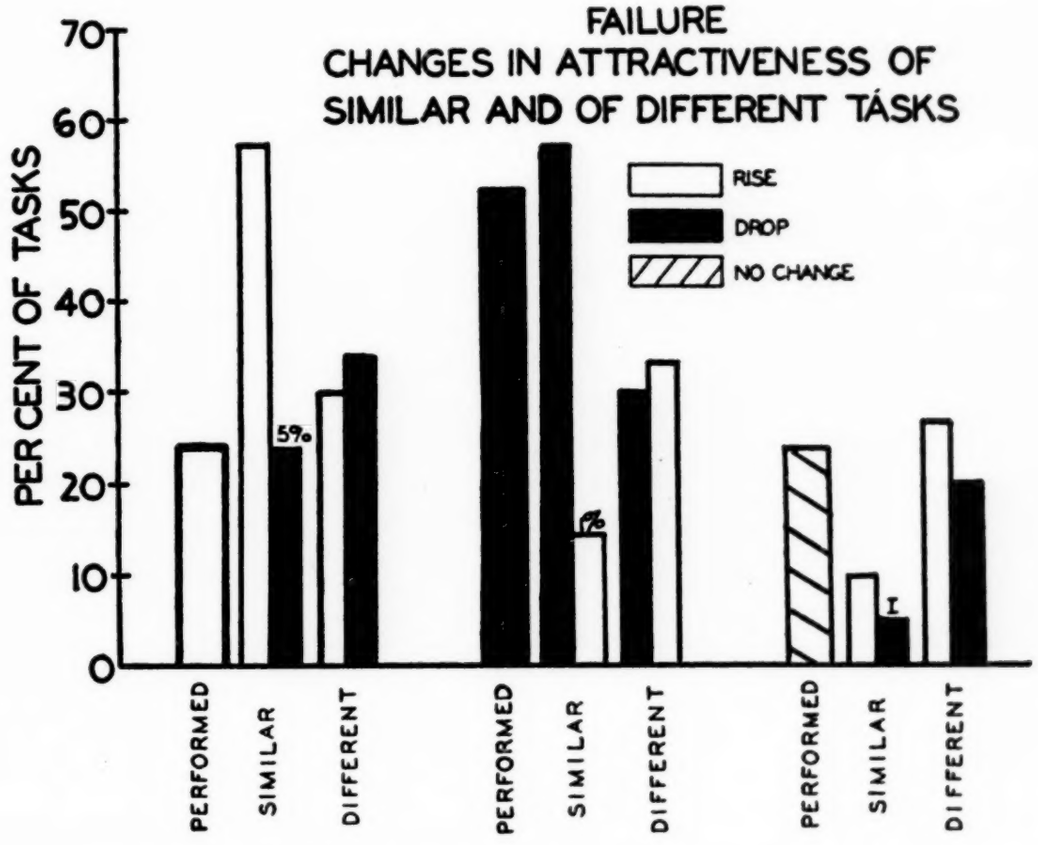

FIG. 4. Percent of similar and different tasks which were raised or lowered on the preference-scale following failure on the performed task.

A fairly sizable group of subjects raised the preference-rating after failure, thus demonstrating the necessity for a theory of failure which can account for both increases and decreases in attractiveness. Interviews with the subjects seemed to indicate clearly that many subjects who raised the preference-rating viewed the experimenter's evaluation of their performance not as failure but rather as an obstacle on their way to the goal. For these subjects failure acted much like interruption. Other subjects seemed to want to 'erase'

the failure by subsequent success on the same activity.[11] For both groups of subjects a tension system could be assumed with its spread to neighboring regions increasing the attractiveness of the similar tasks (Fig. 4).

Rosenzweig has also investigated this problem and has come to rather similar conclusions. His concepts of 'ego-defense' (**21**) and of 'punitive types' (**25, 24, 22**) are designed to account for the different types of effects resulting from failure. They serve the excellent function of emphasizing the influence of certain factors in a situation of failure. The next step forward toward a complete theory of failure seems to be the sharpening of these terms by specifying further their conceptual properties. Until such a step is made it remains difficult to apply them precisely to new experimental situations.

Summary and Discussion

The primary experimental findings may be summarized conveniently under the following eight headings.

1. Interrupted tasks were rated as more attractive after interruption than before by 42 percent of the subjects, less attractive by 22 percent, and equally attractive by 36 percent.

2. Completed tasks were rated as more attractive after completion by 44 percent of the subjects, less attractive by 49 percent, and equally attractive by only 7 percent.

3. After failure 24 percent of the subjects rated the performed task as more attractive, 52 percent as less attractive, and 24 percent left it unchanged.

4. Under all experimental variations, whenever the performed task was raised on the preference-scale, tasks viewed by the subjects as highly similar to it were raised significantly more often than they were lowered; tasks viewed as different from it showed no predominant direction of change.

5. When the completed task was lowered tasks highly similar to it were *raised* significantly more often than they were lowered. There was also a tendency for tasks different from these completed tasks to be raised.

6. When the interrupted task was lowered tasks highly similar to it were lowered more frequently than they were raised; and tasks viewed as different from it were raised more frequently than they were lowered.

7. When the task on which failure was experienced was lowered, tasks similar to it were also lowered more frequently than they were raised.

[11] For this second group of subjects it is possible that one could speak of a 'need for success.'

8. In no case when the performed task was left unchanged was there a significant preponderance of change in either the similar or the different tasks.

These data all combine to support the generalization that perceived similarity is a primary determinant of interdependence. *In every variation of the experiment, when the performed task changed in attractiveness the tasks rated as similar to it changed in attractiveness more frequently and to a greater extent than the tasks rated as different from it.* Two specific types of communication under this general principle have been suggested. The first is the spread of need tension to neighboring systems with similarity as the primary determinant of neighborhood. The second concerns the way in which success or failure on an activity affects the attractiveness of a whole category of activities.

The conception that failure affects categories of activities rather than specific activities has far-reaching consequences. In the first place, it means that failure on one activity will affect the attractiveness of many activities, and thus, a single failure may have extensive effects. Secondly, and more basically, the range and distribution of these effects depend upon the nature and the inclusiveness of the categories which the individual has. For example, the effect of a child's failure on a multiplication test may be to make him dislike multiplication, or arithmetic, or school-work, or even 'everything I do,' depending upon the category to which he subsumes the test.

In conclusion, there is some indication, although the evidence is not statistically conclusive, that anticipation of failure more unequivocally reduces attractiveness than the actual experience of failure. This fact suggests that a person's view of his future is the primarily decisive factor in producing changes in attractiveness. The selection of subjects who *anticipate* failure produces a homogeneous group in respect to the view of the future whereas the selection of subjects who *have* experienced failure does not necessarily do so. Past failure does not always lead to an expectation of future failure.

Appendix

The quantitative data of a study of this kind depend to a rather high degree upon certain technical details. For this reason a description of the more important of these details is here appended.

1. *Selection of Tasks.*—In this experiment tasks were required which could be readily submitted to the three experimental treatments. (1) It was necessary for the experimenter to be able to observe the subject's progress on the task so that it could be interrupted at approximately the same place for all subjects. (2) The same task must allow completion by all subjects within a given range of time. (3) This task must also permit the experimenter to interpret the subject's completion as a failure.

A description of each task is given below. It will be seen that the tasks, upon a basis of the similarity to one another, fall into three groups of four tasks each: peg puzzles, mathematical

problems, and jig-saw puzzles. One task from each group was selected for performance (tasks 1, 5, and 9). The other tasks served simply as indicators of the interactions in attractiveness.

Task 1 is made of a piece of wood ½″ thick with a 10″ × 10″ surface, stained dark brown. On the surface, painted in red, are two concentric squares with the dimensions of 7″ × 7″ and 3½″ × 3½″ respectively. Within the area between the two squares are ninety small holes which permit the insertion of nails. Inserted in these holes are thirty-two nails, grouped in fours so that there are four nails in each corner and four between each corner. The nails are painted red. The problem is to remove four of the nails from the board and to rearrange the remaining nails so that there are still twelve nails on each side of the inner square.

Task 2 is the same as *Task 1* except that the squares are painted black and the nails are left unpainted. The problem here is to take four additional nails and to insert them into the holes, rearranging the nails in such a way that there are still twelve nails on each side of the square.

Task 3 is made of a piece of unpainted wood whose dimensions are 8″ × 10″. An inch inside of the outer edge are drilled twenty-eight holes symmetrically arranged. Eight sticks are inserted in these holes at regular intervals. Two inches inside of the outer row of holes are drilled eighteen more holes forming a rectangle. A stick is inserted in each of these holes. Within this inner rectangle are drilled five holes forming a straight line. No sticks are placed in these holes. The purpose of the task is to rearrange the sticks in eleven moves so that there is a stick in alternate holes all over the board. Although this appears to be a genuine puzzle, its solution is simple.

Task 4, known commercially as Puzzle Peg, is made of a flat square board with a cross drawn upon its surface. Upon the cross are thirty-three squares with a hole drilled in each square. All of the holes, except the center one, contain pegs. The purpose of the task is to remove the pegs by jumping and to jump them in such a way that only one peg is left at the end of the sequence.

Task 5 is typed on a sheet of paper: "A streetcar conductor began his run with no passengers on the car. At the first stop four passengers got on the car and paid cash fares. At the next stop one passenger got off and three times as many passengers as were left on the car got on. All of these paid cash fares. At the next stop six passengers got on the car and all of these gave the conductor transfers. At the next stop one-third of the passengers now on the car who had paid cash fares got off the car. At the next stop twice as many people as were now on the car got on and paid cash fares. At the next stop all of the people who had given the conductor transfers got off the car and one-third of the cash customers did likewise. Finally, thirteen passengers got on the car. Of them, three more passengers paid cash than gave transfers. The fare was seven cents. How much money did the conductor take in during the trip?"

Task 6: "How may a gardener plant eleven trees in such a manner that they shall form twelve rows, with three trees in each row?"

Task 7: "What number which divided by 2 gives a remainder of 1, divided by 3 a remainder of 1, divided by 4 a remainder of 1, divided by 5 a remainder of 1, divided by 6 a remainder of 1, but which divided by 7 gives no remainder?"

Task 8 is made of wood of the same dimensions as the first task. It is stained dark with nine squares painted in red on its surface. Upon each of these squares is placed a white cardboard with a red number on its surface. The cards are numbered from one through nine. Commonly known as the 'magic square problem,' the problem is to arrange the numbers in such a way that the sum of each column, row, and diagonal is equal to fifteen.

Task 9 is a jigsaw puzzle composed of eighteen pieces cut from a piece of three-ply wood whose dimensions are 5″ × 9″. On its surface is pasted a colorful picture of a motion picture actress seated at a table. The pieces are cut irregularly so that their shapes provide cues as well as the picture.

Task 10 is a jigsaw puzzle made of dark brown composition board. Its seven pieces all have straight edges and when fitted properly form a square whose dimensions are 4″ × 4″. Only the shapes of the pieces provide cues for the solution.

Task 11 is made from a piece of three-ply wood whose dimensions are 6″ × 6″. It is composed of ten unevenly shaped pieces whose surfaces are left with the grain of the wood showing. The solution of the puzzle is facilitated by both the grain and the shape of the pieces.

Task 12, a jigsaw puzzle consisting of ten pieces cut from a piece of three-ply wood whose dimensions are 6″ × 6″, has upon its surface a large black paper ring 5″ in diameter and ½″ in width. The pieces are all cut with straight edges so that the grain of the wood and the shape of the paper ring provide the primary cues in the solution.

2. *Schedule of Treatments.*—It appears likely that the temporal order of the presentation of

the experimental treatments is of some importance. The experiment was designed, therefore, so that each of the possible sequences of the three treatments should be given equally often. Further, the preponderant combination of an experimental treatment with a particular task might have biased the results. To avoid this potential disturbance, the experimental design arranged for each of the performed tasks to be combined with each experimental treatment in each temporal sequence equally often. Due to uncontrollable circumstances these requirements were not precisely fulfilled, but they were so closely approximated that little bias could have resulted from this approximation.

3. *Ratings of Similarity.*—At the end of each experimental session three ratings of similarity were made by the subject. Each rating-scale had one of the performed tasks written at the top. The subject was told to place all of the tasks on the scale indicating the degree of similarity to the performed task by the distance on the scale from the performed task. The tasks so clearly formed three natural groupings that all subjects agreed on the groupings. Relative positions within the groups, of course, varied from subject to subject.

(Manuscript received December 15, 1941)

References

1. Adler, D., & Kounin, J. Some factors operating at the moment of resumption of interrupted tasks. *J. Psychol.*, 1939, **7**, 255–267.
2. Barker, R., Dembo, T., & Lewin, K. Frustration and regression: an experiment with young children. *Univ. Iowa Stud. Child Welf.*, 1941, **18**, No. 1.
3. Cartwright, D. Decision-time in relation to the differentiation of the phenomenal field. *Psychol. Rev.*, 1941, **48**, 425–442.
4. Dembo, T. Der Aerger als dynamische Problem. *Psychol. Forsch.*, 1931, **15**, 1–144.
5. Frank, J. D. Recent studies of the level of aspiration. *Psychol. Bull.*, 1941, **38**, 218–226.
6. Freeman, G. L. Changes in tonus during completed and interrupted mental work. *J. gen. Psychol.*, 1930, **4**, 309–334.
7. Hoppe, F. Erfolg und Misserfolg. *Psychol. Forsch.*, 1930, **14**, 1–62.
8. Katz, D. Zur Grundlegung einer Bedürfnispsychologie. *Acta Psychol.*, 1935, **1**, 119–128.
9. Katz, D. *Animals and men.* London: Longmans, Green, & Co., 1937.
10. Katz, E. Some factors affecting resumption of interrupted activities by preschool children. Univ. Minnesota Press, 1938, *Inst. Child Welf. Monog.*, No. 16.
11. Köhler, W. *Dynamics in psychology.* New York: Liveright, 1940.
12. Köhler, W., & von Restorff, H. Zur Theorie der Reproduktion. *Psychol. Forsch.*, 1935, **21**, 56–112.
13. Lewin, K. Formalization and progress in psychology. In *Studies in Topological and Vector Psychology I. Univ. Iowa Stud. Child Welf.*, 1940, **16**, 9–42.
14. Lewin, K. The conceptual representation and the measurement of psychological forces. *Duke Univ. Series, Contr. Psychol. Theor.*, 1938, **1**, No. 4.
15. Lissner, K. Die Entspannung von Bedürfnissen durch Ersatzhandlungen. *Psychol. Forsch.*, 1933, **18**, 218–250.
16. Mahler, W. Ersatzhandlungen verschiedenen Realitätsgrades. *Psychol. Forsch.*, 1933, **18**, 27–89.
17. Marrow, A. J. Goal tensions and recall. (I & II), *J. gen. Psychol.*, 1938, **19**, 3–35; 37–64.
18. Nowlis, H. H. The influence of success and failure on the resumption of an interrupted task. *J. exp. Psychol.*, 1941, **28**, 304–325.
19. Ovsiankina, M. Die Wiederaufnahme unterbrochener Handlungen. *Psychol. Forsch.*, 1928, **11**, 302–379.
20. Pachauri, A. R. A study of Gestalt problems in completed and interrupted tasks. *Brit. J. Psychol.*, 1935, **25**, 447–457.
21. Rosenzweig, S. Need-persistive and ego-defensive reactions to frustration as demonstrated by an experiment on repression. *Psychol. Rev.*, 1941, **48**, 347–349.
22. Rosenzweig, S. The experimental measurement of types of reaction to frustration. *In* Murray, H. A., *Explorations in Personality.* New York: Oxford Univ. Press, 1938.
23. Rosenzweig, S. Preferences in the repetition of successful and unsuccessful activities as a function of age and personality. *J. genet. Psychol.*, 1933, **42**, 423–441.

24. Rosenzweig, S. A test for types of reaction to frustration. *Amer. J. Orthopsychiat.*, 1935, **5**, 395–403.
25. Rosenzweig, S. Types of reaction to frustration. *J. abnorm. (soc.) Psychol.*, 1934, **29**, 298–300.
26. Sandvoss, H. Ueber die Beziehungen von Determination und Bewusstsein bei der Realisierung unerledigter Tätigkeiten. *Arch. ges. Psychol.*, 1933, **89**, 139–192.
27. Schlote, W. Ueber die Bevorzugung unvollendeter Handlungen. *Z. Psychol.*, 1930, **117**, 1–72.
28. Sliosberg, S. Zur Dynamik des Ersatzes in Spiel und Ernstsituationen. *Psychol. Forsch.*, 1934, **19**, 122–181.
29. Zeigarnik, B. Ueber das Behalten von erledigten und unerledigten Handlungen. *Psychol. Forsch.*, 1927, **9**, 1–85.

INTRASERIAL INHIBITION AS MEASURED BY REPRODUCTION

BY HELEN E. PEIXOTTO

Fordham University

Inhibition of memory as an explanatory hypothesis for forgetting has been investigated by two methods. In one method the activity of interpolated material causes the inhibition—this is the type which Foucault (3) describes as 'external inhibition'; in the other method the inhibitory activity has its origin in the interior of the material to be recalled and is, therefore, a function of the original learning situation, termed by Foucault 'internal inhibition.'

Distinction has also been made between retroactive and proactive inhibition; or, in Foucault's terminology, 'regressive' and 'progressive' inhibition. Retroactive inhibition is the 'detrimental influence of subsequent activity upon the retention of previously established activities' (1). Proactive inhibition may be defined conversely as the detrimental influence of prior activity upon the retention of subsequently established activities.

This experiment has been devised to investigate intraserial retroaction and proaction in terms of reproduction, or in other words to study the phenomena of retroactive inhibition and proactive inhibition as functions of serial position and as measured by reproduction.

Historical Survey

Experiments on retroactive inhibition have covered many phases of the problem. In this investigation, however, we are primarily interested in inhibition as a function of serial position and therefore shall make reference only to those experiments directly concerned with serial position. It is interesting to note that in Britt's review of the literature through 1935 (1), the only experiment definitely related to this topic is that of Foucault cited above. However, there are several recent experiments on serial position that must be considered although they differ markedly from the present experiment, chiefly in their use of the method of interpolated material.

It might be well to describe here in more detail the techniques used to elicit internal inhibition. In the method used in this study the material to be learned is necessarily a continuous series wherein each item or items act as interpolated material on those preceding

them. Thus, the interpolated material is contained in the original learning situation.

In two experiments by McGeoch and McGeoch (**6**) and McGeoch and Sisson (**8**) on serial position an attempt was made to change the order of presentation but no significant difference in inhibition was found to accompany change of position. The method of rotation was, however, inadequate because, while there was a change in the absolute position within a series, intraserial sequence remained the same. The conclusions of these two experiments are contradictory. This contradiction the authors say may be explained by length of list being a factor in inhibition as a function of serial position.

In an experiment performed later by McGeoch and McKinney (**7**) on order of presentation it was found that the amount learned is a function of the constancy of the order of the pairs at successive presentations. When the order, that is, the serial positions of the pairs, is the same throughout, more is learned than when the order is different at each presentation. This is true whether the order in which the pairs follow each other, regardless of serial position, is kept the same or is varied randomly. Amount of inhibition is a function of the relative order of the pairs in the original and interpolated lists. McGeoch and McKinney conclude that these results are not primarily a function of the differences in degrees of learning under the two types of presentation, nor are they a function of the intraserial order of items as separate from their serial positions.

Sand (**14**), in an experiment on retroactive inhibition, makes use of the usual method of investigation, that of interpolated material. The conclusions derived from this investigation are that lists of paired nonsense syllables are relatively more difficult to learn than short lists; decreasing susceptibility to retroactive inhibition with increasing length of list is a result of the greater overlearning which is given longer lists; serial position does not influence the difficulty of paired nonsense syllables (as they were learned in this experiment) as much as it influences the difficulty of syllables of a list which is learned as a whole.

Mitchell (**10**, **11**) and Schmeidler (**15**) have also performed experiments related to the present investigation, but in both instances the number of subjects was inadequate. Mitchell, using numbers, found a 'form of retroactive inhibition within the list.' This inhibition was not found by the technique of interpolated material. Schmeidler, in the treatment of her results, discusses both retroaction and proaction, but her data are suggestive rather than conclusive.

Experiments in which the material used was geometric forms constitute another type of investigation within the field of inhibition. Despite the different form of material employed, the experiments of

Gibson and Raffel (4), Crafts (2), and Philip (12) are intimately connected with the present one because the method of interpolated material is not used. Even among these investigations, however, results are contradictory. Crafts finds primacy and finality sufficient to explain his results; Gibson and Raffel and Philip find that this law is inadequate to account for their results. Gibson and Raffel conclude that in their experiment retention was an inverse function of serial position when five forms were given, but did not vary with position when only three forms were given. They state that there is no effect of primacy as such in these results; and that recency has been shown to depend principally on subsequency of serial position rather than on recency as such. The results of Philip can be partially explained in terms of proactive and retroactive inhibition. He also found serial position to be unimportant in short lists, but that "a series of impressions is different in its total effect when the series bears the character of group membership;" and that proaction may be as important as retroaction in a short series.

The experiment of Robinson and Brown (13) on the effect of serial position is not an experiment directed towards the study of inhibition, but must be considered in any investigation on this subject. The conclusions of this experiment indicate the advantages for recall of primacy and finality, the curves of memorization being negatively accelerated. The deceleration of the learning rate varies directly with the serial position of the item, and there is positive acceleration in learning for the items unfavorably situated in very long lists.

The experiment most intimately connected with the present one is that of Foucault (3). Foucault used lists of meaningful words presented serially and which his subjects, one hundred children, attempted to reproduce immediately. The results were discussed along the following lines: regressive inhibition as measured by failures of the first and second words of each series, progressive inhibition as measured by failures for the last and next to the last words of each series, and progress of failures of each series of increasing length, that is, percentage of failures, which is a measure of the total internal inhibitory processes.

As measured by failures for the first and second words of each series, the results indicate a linear relationship between serial position and regressive inhibition. Progressive inhibition as measured by failures for the last and next to the last words of each series yields an irregular curve similar to that for regressive inhibition.

Concerning these results, Foucault says (p. 103): "Regressive and progressive inhibition also act on the second and the next to the last term of a series. . . . The second term, at the moment when it

commences to undergo the regressive inhibition of those which follow it, has already undergone progressive inhibition excercised by the first. Consequently the action of progressive inhibition consists in stopping the psychological development of the word on which it is exercised. . . . On this hypothesis is explained the increasing failures for successive terms in the series as one proceeds towards the middle. In the second half of the series the number of failures decreases as one approaches the last term, this is explained by the diminution of regressive inhibition which is zero for the last term." The present experiment was undertaken in further substantiation of this hypothesis.

Procedure

A. Methodology

The technique employed in this experiment is based on the following argument of Foucault: If a subject is able to learn and recite a series without error we may say there is no knowable action of internal inhibition. But if the series is lengthened, a point should be reached where two kinds of inhibition are manifest: regressive inhibition will be manifest by failures at the beginning of the series; and progressive inhibition by failures at the end of the series (3).

In order that these phenomena of inhibition should appear to the highest possible degree, it was decided to use paired associates of nonsense syllables. Nonsense syllables were chosen to overcome the primary weakness of Foucault's investigation, which was the high associative value of the material. Moreover, due to their low associative value, they limit learning to that which takes place in the experimental situation.

The method of paired associates was decided upon because in this method an item in question is voluntarily associated only with its paired associate and not with the items preceding and following it. To state the idea differently, in this method we are dealing with 'horizontal' association which forms bonds between the pairs of associates. Opposed to this is 'vertical' association which deals with the serial linkage between items, and which occurs in the anticipation method.

Considering these factors it is obvious that the anticipation method is ill-adapted to bring out inhibitory phenomena. In this procedure the subject intentionally attempts to associate each item with the succeeding one. He attempts to create and strengthen the associative mechanism so that each item will evoke the following one.

On the other hand the method of paired associates is intrinsically adapted to the demonstration of inhibitive effects. In this method the subject attempts to form associations, to be sure. But these associations are between the first and the second members of each pair, not between the serial items. As a matter of fact, the concentration upon the formation of associations between the members of each pair minimizes the formation of associations between the serial items. This is particularly true where, as in this experiment, the items in the recitation series are completely rotated. Under such conditions, attempts by the subject to learn the serial position of the pairs as well as to establish linkage between the two members of the pairs will definitely interfere with his performance in the recitation.

By the anticipation method, where the learning order and the recitation order must perforce be the same, vertical association alone operates and the serial relationship is stressed. By the paired associate method, where the recitation order is varied, though the learning order is maintained constant, both vertical and horizontal associations come into play. From the former method we can gauge the difficulty of memorization of the various items in the series, chiefly by facility in producing the sequence; the latter method enables us to measure serial difficulty by the accuracy of reproduction of the second members of pairs when the first members are given. The anticipation method would naturally stress the first members of the series since they start the sequence; one would not expect as much stress on the initial members of the series in the paired associate method where the recitation order is rotated.

In the present experiment inhibitory influences of intraserial factors were studied rather than

inhibition due to interpolated materials. In the light of previous experiments this method should show larger amounts of inhibition than the method of interpolated material. Britt (1), in reviewing the literature, found results were contradictory concerning the effects of temporal position. However, for very short periods, it is agreed that inhibition is greatest when the interfering activity immediately precedes recall or immediately follows learning.

Moreover certain previous studies have shown that inhibition varies directly with the similarity of the material. In this study we attempted to achieve the highest degree of similarity without identity. A review of the literature on this topic may also be found in Britt's article.

Learning was measured by reproduction. The methods of recognition and reproduction were the only ones which could be used in group experimentation. Recognition has not been found to be very satisfactory in experiments on inhibition since the amount of inhibition is relatively small; therefore the method of reproduction was decided upon (12).

B. Preliminary Experiments

Three preliminary experiments were conducted. In all of these nonsense syllables from the Glaze (5) list of syllables having zero percent associative value were used; the time intervals employed were two seconds for the learning presentation and five seconds for the recitation presentation. In the first preliminary experiment the syllables were presented tachistoscopically, *using the same serial order in both the stimulus and response lists.* The second preliminary experiment differed from the first in that it was a group experiment presented by means of motion picture film. In this experiment *there was no rotation of syllables in the recitation series, but the order of the syllables was different from that of the learning series.* The third of these experiments was conducted to check the associative value of the Glaze syllables. Ninety-nine syllables were studied for the qualitative and quantitative value of each syllable. From this list thirty-two syllables were chosen, sixteen from the 6.67 percent list of Glaze and sixteen from the zero percent list.

The results of these experiments indicated that a three syllable list was too easy to show inhibitory effects, and that complete rotation in the recitation series was necessary. The learning curve for the second preliminary experiment indicated a shift in the learning process. That is, owing to the fact that a constant learning order and a constant but different recitation order were used, somewhere in the course of the learning the subjects shifted their method of learning from the order of the learning series to that of the recitation series.

C. Major Experiment

The procedure for the major experiment was developed from the results of the preliminary experiments. Approximately six hundred high school students participated as subjects, of whom five hundred and four gave complete records. The procedure consisted of the presentation of a film on which the syllables appeared in white letters on a black ground. By this procedure the time intervals were well controlled. The time intervals in the presentation were two seconds for each pair in the learning series, and five seconds for each stimulus syllable in the recitation series to give the subject time to write the response before the next syllable appeared on the screen. There was a two second pause between the last pair of syllables of the learning list and first syllable of the recitation list, and likewise between the last syllable of the recitation list and the first pair of the learning list, when there appeared on the screen the words 'Ready to Write' and 'Turn Your Page' respectively.

Five series of syllables were used, consisting of lists of four, six, eight, ten or twelve paired associates. These paired associates appeared on the screen one pair at a time in the learning series. In the recitation series the first syllable only of each pair appeared, only one syllable being on the screen at a time. In the recitation series the subjects were required to write the second syllable of each pair, that is the syllable that had appeared with the given stimulus syllable in the learning series. By presenting each pair one at a time in the learning series the amount of time spent on learning was similarly limited for each subject. After each learning series a recitation series was presented.

Sixteen pairs of syllables were used in all. This had the advantage of allowing for rotation of the syllables interserially so that any increased facility in recall could be determined from position, independent of the associative value of the individual syllables. There was no rotation within a learning series since this would tend to neutralize the effect of serial position, but the same pairs

were used from list to list, no pair occuring twice in the same relative position. Each pair appeared in a relatively weak and strong position in at least one list. The strong positions were the terminal ones, and weak positions were the center ones of the series, in conformity with the findings of Robinson and Brown (13). The 4-Series was given first to each group of subjects that took the other series, and was composed entirely of different syllables from those of the other lists. The syllables in the 6-Series list appeared in every possible relative position. Different groups of subjects learned each of the 6-, 8-, 10-, and 12-Series.

The stimulus syllables of the recitation series were in different order for each presentation and were completely rotated. This rotation was undertaken to insure that any variation in facility of learning would be dependent on serial position in the learning series rather than on factors extraneous to position. It also eliminated the possibility of learning the response list serially in the learning period, instead of learning to respond to the first element of the pairs by giving the second member. This precaution was found to be necessary from the results of the first preliminary experiment.

The number of presentations per series was one and one-half times the length of the series; thus the 6-Series was presented nine times, the 10-Series, fifteen times. This number was determined by preliminary experimentation as being the least number of trials required to produce sufficient learning by the average subject. (Complete learning by all subjects was not desired.)

TABLE I

SCHEME OF ROTATION OF THE SYLLABLES IN LEARNING ORDER FOR EACH SERIES

6-Series	8-Series	10-Series	12-Series
1. ZOF-GEX	YIX-ZOJ	ZUV-CIJ	QOS-MEF
2. YAV-WUQ	ZUV-CIJ	VUF-ZEQ	KOJ-YUF
3. GEC-VAF	MIB-QUJ	YIF-VUQ	QIM-JEX
4. ZUV-CIJ	QOS-MEF	YAV-WUQ	GEC-VAF
5. QOS-MEF	YAV-WUQ	YIX-ZOJ	ZOF-GEX
6. YIX-ZOJ	VUF-ZEQ	ZOF-GEX	MIB-QUJ
7.	GEC-VAF	QOS-MEF	VUF-ZEQ
8.	ZOF-GEX	QIM-JEX	YIX-ZOJ
9.		MIB-QUJ	ZUV-CIJ
10.		GEC-VAF	YIF-VUQ
11.			QOJ-CEF
12.			YAV-WUQ

Before the experiment a printed pad containing a separate page for each recitation was given to each subject and then the experimenter gave the instructions orally. An assistant operated the movie projector. The experimenter, the assistant and the teacher of the class assisted in proctoring.

Preliminary instructions were given before the 4-Series; between the 4-Series and the one following there was a four minute interval. During this period questions were encouraged in order to clarify the instructions. The rest period also served to eliminate transfer effects from the 4-Series to the one following it. There was also extra emphasis put on certain phases of the instructions and remarks of encouragement were given during this interval. Instructions were as follows:

> "This is a memory test. The material to be remembered is nonsense syllables, that is, combinations of letters that do not make words. You will be shown four pairs of syllables. As they appear pronounce them to yourself and try to remember them. For instance, K-E-Z—T-O-R would be KEZ-TOR. Then you will be shown only the first syllable of each pair. As each of these first syllables appears write on the pad the syllable that went with it. Write only one syllable, the second syllable. This task is not easy and you probably won't get them all the first time. So we will repeat the whole procedure several times. Try as hard as you can all the time. Look at the pads. Fill in the cover page. Now turn the page. On the first page it says Recitation 1 and you see the numbers 1 to 4. As the first single syllable appears write opposite number 1 the syllable that went with it. When the

second syllable appears write opposite number 2 the syllable that went with it. And so on. If you do not know the syllable to write down, draw a line so that you will get the next one in the right space. Use the first page for the first trial, the second page for the second trial and so on. It will say on the screen when you are to turn the page. Turn it quickly when it says 'Turn your page' because you are given only two seconds. Do not write until it says on the screen 'Ready to write.' Are there any questions? Silence must be kept during the test."

If there were any questions the directions were repeated until the class said they understood what they were to do. Then the experimenter said 'Ready' and test was begun.

When the four syllable list was finished the projector was stopped. The class was told to put their pencils down. They were asked if there were any questions; if so, these questions were answered. They were then told that the next part of the test was done in exactly the same way except that it was longer. They were reminded to keep their pencils up when they were not writing and to draw a line when they did not know the response. They were also encouraged to continue trying even if the task seemed very difficult at first. The experimenter then said 'Ready,' and the next series was presented. The subjects were tested as class units in their usual class room situation. Table II reports the number of subjects of each sex participating in each series.

TABLE II

Series		Number	Girls	Boys
4	6	129	95	34
4	8	115	95	20
4	10	139	48	91
4	12	121	42	79
Total	4	504	280	234

Results and Interpretation

A. Scoring

The records obtained from each subject were scored according to the following code:

A—Correct structure, correct position
B—Correct structure, one position removed
C—Correct structure, more than one position removed
D—Reversal in structure, correct position
E—One error in structure, correct position
F—Reversal in structure, incorrect position
G—One error in structure, incorrect position
X—Entirely wrong response
O—No response

In order to ascertain the effect of serial position, which differed for learning and for recitation, the scores based on the above code were tabulated first, according to recitation order, and secondly, according to learning order. The recitation order tabulations gave the preferred position in recall; the learning order gave the preferred positions in the learning.

It was found that the E scores gave the same curve as the A scores though with diminished frequency, and that the B, C, G, and X scores did not appreciably differ with position. D and F scores were very infrequent. Hence it was decided for the purpose of this experiment to use only the A scores on the major series following the general practice of workers in this field.

The total scores for each position in the recitation order and for each syllable in the learning order were averaged for the number of subjects who took each series and for the number of trials in each series. This is referred to as A.S.T. (Average per Subject per Trial).

B. Comparability of the Groups

The four groups used in this experiment were chosen at random from among twenty-four high school classes. It was necessary to know if these groups were equated in order to justify a comparison of the results of the various series. Moreover, comparability of the groups in this particular type of task was desirable rather than comparability in any extraneous abilities such as intelligence, or unrelated factors such as age, or grade placement.

It will be recalled all experimental groups were given two series—the 4-Series and one of the remaining series (6, 8, 10, 12). The learning and recitation orders for the 4-Series were exactly the same for each experimental group, thus making the results on the 4-Series a measure of the comparability of the groups.

This analysis is presented in Table III. The critical ratios of the differences in means were determined among the various groups according to the results of the 4-Series. The formula Diff./σ diff. was used. Differences are considered to be significant if the critical ratio is 3.00. As may be observed from Table III none of the critical ratios reach this criterion of significance. Therefore in regard to performance in this experiment the groups were presumably comparable on the six, eight, ten and twelve item tests.

TABLE III

CRITICAL RATIOS ON THE 4-SERIES

(A and B Scores)

Series	C. R.	Series	C. R.	Series	C. R.
6 and 8	1.070				
6 and 10	.468	8 and 10	.616		
6 and 12	1.682	8 and 12	.518	10 and 12	1.182
		(A Scores)			
6 and 8	1.780				
6 and 10	2.880	8 and 10	1.039		
6 and 12	.998	8 and 12	.919	10 and 12	1.960

In both A, and A and B Scores all the critical ratios are below 2.00 except in one case in the A Scores where it is 2.8, which is, however, still below 3.00.

C. Raw Scores

The supremacy of certain syllables in ease of learning is constant throughout the various series. That is, in general, when the trials are grouped in threes the relative position of each syllable in regard to its learning score is maintained. These data are shown in Table IV. The data in this table are based on raw scores, the totals not having been treated for differences in number of subjects per group, nor for differences in trials per series. This correction was not necessary here since this table is only for the purpose of intraserial comparison, not interserial comparisons. From Table IV the method of learning of the subjects can be seen. It can be seen that certain syllables act as nuclei for the spread of learning. Thus syllable 5—6-Series (QOS-MEF); syllable 4—8-Series (QOS-MEF); syllable 8—10-Series (QIM-JEX); syllable 3—12-Series (QIM-JEX); from the first trials to the last are generally the easiest syllables, and syllables adjacent to those mentioned are, in general, more easily learned than syllables more remote. Thus the aforementioned syllables maintain their supremacy in ease of learning and they served as foci or points of departure for the acquisition of other syllables within the list, in other words certain syllables act as nuclei for the spread of learning.

A probable reason for this type of learning is the extreme difficulty of the learning situation, which was necessary to bring out the inhibitory processes. This difficulty was obtained by means of five factors:

1. Short exposure time of each pair of syllables in the learning presentation.
2. Choice of syllables of zero percent associative value according to Glaze. Choosing sylla-

bles of such low value made them of unusual pronunciation which added to their inherent difficulty. Besides there is naturally more deviation among the syllables of this group—those of zero percent associative value—since they are at the end of the Glaze distribution, than there is among syllables nearer to the middle of the distribution.

TABLE IV

TOTAL NUMBER OF CORRECT RESPONSES

(A Scores totaled for each group of three trials)

Series	Groups of 3 Trials	Order of Syllables											
		1	2	3	4	5	6	7	8	9	10	11	12
4	I	7.25	11	32	36								
	II	15	22.5	70	72.5								
6	I	12	2	11	11	72	19						
	II	32	6	35	11	109	16						
	III	5	25	27	19	131	32						
8	I	2	9	15	34	1	1	8	33				
	II	1	11	44	64	5	10	18	44				
	III	4	16	58	98	15	7	33	29				
	IV	11	8	66	149	13	7	52	75				
10	I	5	0	1	4	2	5	13	46	5	13		
	II	9	3	2	7	2	9	41	33	31	93		
	III	12	7	6	8	7	17	62	50	33	59		
	IV	20	8	12	17	10	32	63	113	54	142		
	V	31	11	11	17	14	34	83	104	79	110		
12	I	21	6	40	3	2	7	0	5	0	0	7	3
	II	29	8	51	19	4	13	5	0	3	3	20	12
	III	52	8	66	34	13	19	4	3	7	6	15	5
	IV	77	8	76	59	14	42	13	3	12	6	25	16
	V	99	16	75	88	20	58	6	7	16	4	42	15
	VI	87	28	73	99	31	86	17	11	22	7	62	10

In order to compare the progress of the learning within the series the total number of A scores for each position of each learning series is given in this Table for each group of three trials. The total scores for the 4-Series were divided by four to make these scores more comparable with those of the other series. This was necessary since the 4-Series was given to all the groups and therefore had approximately four times as many subjects as did the other series. It must be remembered, however, that these raw scores are not exactly comparable from series to series because of the varying number of subjects and trials for the different series. Interserial comparability is achieved by the A.S.T. score, which appears later.

3. Many of the syllables were of similar pronunciation and spelling which caused confusion. Examples are GEX and JEX, MEF and CEF. Other examples may be seen in Table I.

These three factors resulted in the tendency for one syllable to act as a point of departure from which the learning spread to other syllables. The fact that the same syllables acted as nuclei for learning for the entire group to which a series was given, indicates that the explanations for this mode of the spread of learning are inherent in the syllables themselves.

4. Besides the difficulties inherent in the syllables, the methodology of the experiment was so devised as to make learning difficult as has been previously discussed. Thus the use of paired associates instead of a single syllable demanded the setting up of both horizontal and vertical associations.

5. The factor of rotation in the recitation order also influenced the learning situation. Thus the subjects were not permitted merely to learn a single list or a double list, but were forced to form associations presented in the learning order and to respond to the cues presented in the recitation order.

However, no matter how paltry the amount of learning consequent upon the influence of these factors, militating as they do against a high degree of learning, the general trend of the curves is the same. This is true both intraserially and among the various series.

D. Associative Value of Syllables

The irregularities of the curves derived from the raw data made it necessary to analyze the results in terms of separate syllables to determine whether the irregularities were inherent in the learning situation or were caused by certain extraneous factors. This analysis is shown in Table V. The last column of this table gives the averages per subject per trial per list, which is

TABLE V

ASSOCIATIVE VALUE IN RECALL FOR RESPONSE SYLLABLES

Syllables	List	Position	Total Score	A.S.T.	A.S.T. per List
1. QOS-MEF	6	5	350	.2842	.1796
	8	4	335	.2420	
	10	7	261	.1251	
	12	1	364	.1671	
2. QIM-JEX	10	8	346	.1659	.1765
	12	3	408	.1872	
3. GEC-VAF	6	3	70	.0602	.1179
	8	7	113	.0816	
	10	10	403	.1932	
	12	4	298	.1367	
4. MIB-QUJ	8	3	186	.1347	.1132
	10	9	209	.1002	
	12	6	228	.1367	
5. QOJ-CEF	12	11	174	.0797	.0797
6. ZOF-GEX	6	1	40	.0535	.0640
	8	8	187	.1355	
	10	6	99	.0474	
	12	5	87	.0396	
7. KOJ-YUF	12	2	80	.0367	.0367
8. ZUV-CIJ	6	4	46	.0394	.0392
	8	2	39	.0282	
	10	1	78	.0374	
	12	9	59	.0267	
9. YAV-WUQ	6	2	36	.0310	.0282
	8	5	36	.0260	
	10	4	53	.0258	
	12	12	65	.0298	
10. YIX-ZOJ	6	6	74	.0636	.0264
	8	1	20	.0144	
	10	5	33	.0138	
	12	8	30	.0137	
11. VUF-ZEQ	8	6	21	.0152	.0161
	10	2	29	.0138	
	12	7	42	.0192	
12. YIF-VUQ	10	3	34	.0163	.0143
	12	10	27	.0125	

used as a measure of associative value in recall. These values were found by dividing the total score, that is the total number of correct responses for each syllable, by the number of subjects to whom that syllable had been presented, the trials the syllable had had, and the number of lists in which it had appeared. From this table, it is obvious that factors extraneous to serial position were operating, and that all the syllables were not of equal difficulty. That these factors are other than serial position is evident because the final average values for each syllable (associative value in recall) found in column 6 do not differ much from those values for the same syllable obtained in each series (average per subject per trial or A.S.T. values) found in column 5, in spite of the rigorous method of rotation of the syllables shown in this Table and more clearly in Table 1. Although the number of presentations did not permit complete rotation the effect of positions did average out.

It is shown, therefore, that although the syllables are of zero percent associative value according to Glaze, associative value and associative value in recall are not the same, and associative value in recall is quite different for various syllables. The reasons for these discrepancies are discussed in detail elsewhere; they are, namely, association effect, ease of pronunciation and confusion caused by the occurrence of common letters.

Using the Spearman rank order correlation method the associative values found in the preliminary experiment correlate .55 with the associative value in recall. It can be stated, therefore, that the recall value of a syllable is determined more by other factors than it is by free association value.

That associative value in recall is an important factor can be seen in column 5 of Table V. The consistency of the values from the various syllables in each series is very high except for items 1, 3, 6 and 10 where there is one marked inconsistency in each. It is readily observed, however, that each of these inconsistencies occurs when the syllable in question appears in a key position, that is, in a position greatly influenced by serial order. In items 3, 6 and 10 this position is the last in the series and therefore the recall values are enhanced; in item 1 the position is seventh in the 12-Series, a weak position in serial order and the value is accordingly low.

The syllables used in the 4-Series do not appear in this Table because they were not used in any other of the series. Since they always occupied the same serial position, no averages independent of position could be obtained for them.

In view of these differences in difficulty of recalling various syllables having comparable associative values, a scheme of compensation or weighting had to be devised to make all the syllables of equal associative value in recall in order to indicate the effects of serial position as such. The values for the 6-, 8-, 10- and 12-Series are based on the A.S.T. values and corrected for the associative value in recall for each syllable. The method by which this was done varies for the learning and recitation orders so each will be described separately. The method for finding the associative value independent of recall has already been described.

The method of correcting for varying association values of syllables by weighting is as follows:

Recitation Order

A fraction was formed, the reciprocal of which was the corrective value, as follows:

1. Determine the number of times each syllable appeared in the first recitation position, the second recitation position etc.

2. For each position in recitation order, (*A*) multiply the associative value in recall of each syllable appearing in the recitation position by the number of times it appeared in that position; and (*B*) determine the average of these group values.

3. Average the summation of the totals for the several positions.

4. Weight the totals for each position by the averages found in (3).

5. Weight the A.S.T. for each position by the percent averages for each position found in (4) thus obtaining the corrected value.

Learning Order

A corrective factor was obtained similar to that for the recitation order with the following variations:

1. Find the average of the associative values in recall for the syllables of the list.

2. Weight the associative value in recall for each syllable by the average obtained in (1).

3. Weight the A.S.T. for each syllable by the percent average for the corresponding syllable as obtained in (2). This gives the corrected value for each syllable.

E. Discussion

The inhibitory effect on serial position, which is the problem of primary interest to this experiment, is shown by an analysis of the number of correct responses for each syllable in the learning order and for each position in the recitation order. This analysis is presented in Tables VI and VII and the accompanying Figures. Also

TABLE VI

Corrected Values for Syllables for Each Series

Learning Order

	6-Series	8-Series	10-Series	12-Series
1	.0391	.0395	.0873	.0687
2	.0822	.0620	.0658	.0769
3	.0381	.0858	.0875	.0782
4	.0895	.0974	.0702	.0855
5	.1184	.0667	.0401	.0459
6	.1301	.0682	.0569	.0681
7		.0501	.0534	.0880
8		.1620	.0723	.0382
9			.0679	.0603
10			.1257	.0634
11				.0735
12				.0780

TABLE VII

Corrected Values for Syllables for Each Series

Recitation Order

	6-Series	8-Series	10-Series	12-Series
1	.1372	.1456	.1328	.1151
2	.0823	.0822	.0772	.0827
3	.0940	.0650	.0934	.0864
4	.0782	.0802	.0629	.0816
5	.0685	.0788	.0602	.0768
6	.0568	.0450	.0592	.0778
7		.0626	.0538	.0488
8		.0772	.0467	.0551
9			.0621	.0564
10			.0578	.0501
11				.0621
12				.0690

this analysis is discussed in terms of inhibition as related to the aforementioned experiment of Foucault.

Regressive inhibition as expressed by Foucault, quoted herein, that is, 'the inhibitory effect of the consequent image on the antecedent image,' is definitely shown in the results of this experiment. This regressive inhibition is shown very clearly in Fig. 1, the last

item within each series showing an appreciably higher score than any of the other positions. The effect, therefore, of the consequent 'images' on the antecedent 'image' is much stronger than the effect of the antecedent 'image' on the consequent 'image.'

These results are entirely in accordance with those of Foucault, and indicate the more clear cut, stronger influence of regressive inhibition over progressive inhibition.

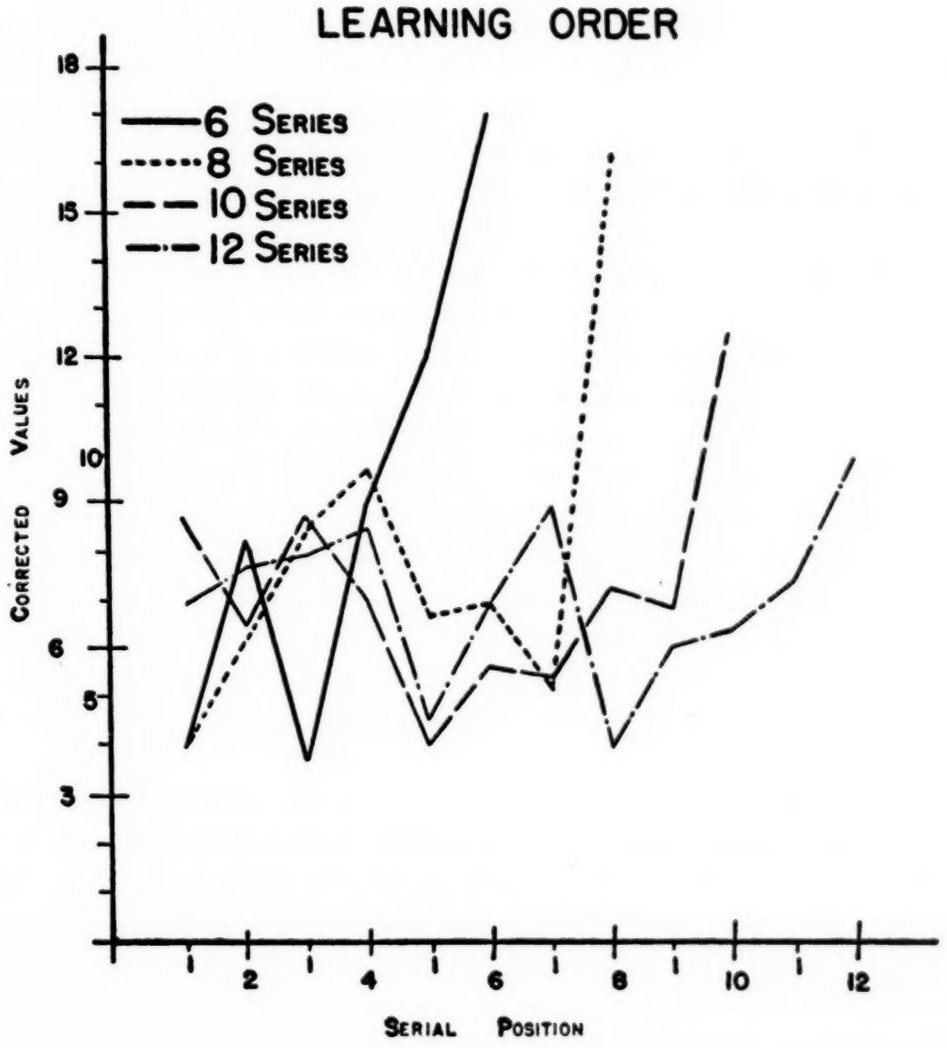

FIG. 1. Variation of difficulty of learning according to serial position for the four series. Difficulty is measured by the average correct response, per subject per trial per list after the syllables are weighted for differences in association value. The learning order of the syllables is constant in each series, but varies from series to series.

These phenomena of regressive and progressive inhibition are not in accord with those investigations on primacy and finality or recency where primacy was found in each instance to be more advantageous than finality. There is one exception among these investigations, that of Welch and Burnett (**16**). They found finality to be the most important factor in learning items in serial order.

Their results are, however, an artifact of their experimental procedure. They presented nonsense syllables tachistoscopically and immediately after the presentation had the subjects write down the syllables they remembered. Naturally they recorded first those seen last. Such an artifact of experimental procedure influencing the results of this experiment is impossible because of the rotation in the presentation of recitation series.

It is true that the first few learning order positions have an advantage over the central learning order positions. This advantage for the first part of a series is similar to the primacy factor found by Robinson and Brown and others. However, it is in no instance in our experiment as high nor as advantageous as these investigators have claimed. We would say, therefore, that this initial rise in the curve is not the same phenomenon as that of primacy described in the aforementioned experiments since, although a tendency to the same type of curve is present, it does not have the same characteristics as other primacy curves, that is a definitely supreme advantage over any other position within a series. It is, however, in accordance with curves of inhibition as found by Melton, Philip and others.

It will be noted that the curves presented in Fig. 1 all show, except Series 10, an initial deflection in the first position. This is probably caused by the short interval of two seconds between recitation and learning among the various trials. It is apparent this interval was too short so that the subjects were not ready for the first pair of syllables when the learning order was presented for a new trial, thereby often missing or failing to get an adequate impression of this first syllable.

Any comprehensive explanation of the phenomena appearing in this experiment must consider the curves shown in both Figs. 1 and 2. If we restrict our attention to the curve of learning order, relative finality would seem to be the condition most conducive to recall. This in turn would, in terms of inhibiting influence, find its explanation in the decreasing amounts of regressive inhibition as we move from the first to the last position in the series; that is, the amount of material interpolated between learning and recitation is progressively decreased. Primacy is of little moment and proactive effects are not readily discernible except in the comparison of the curves for different series, particularly for the last position in each series.

Interserial proactive inhibition is discussed by Foucault in terms of 'internal progressive inhibition.' This inhibition is shown in this experiment in the curve for the learning order, Fig. 1, by the successively lower values for the last syllable as one proceeds through the longer series. The same interserial inhibition is evident in the curves for the recitation order, Fig. 2, except for a slight inversion between

Series 6 and 8. Therefore Foucault's concept of proactive inhibition, that is, increased inhibition depending on the increased length of the series, is herein evidenced since the value of the first syllable in this curve decreases as the amount of preceding learning increases, and is evidenced in Fig. 1 by the value of the last syllable decreasing as the amount of learning increases.

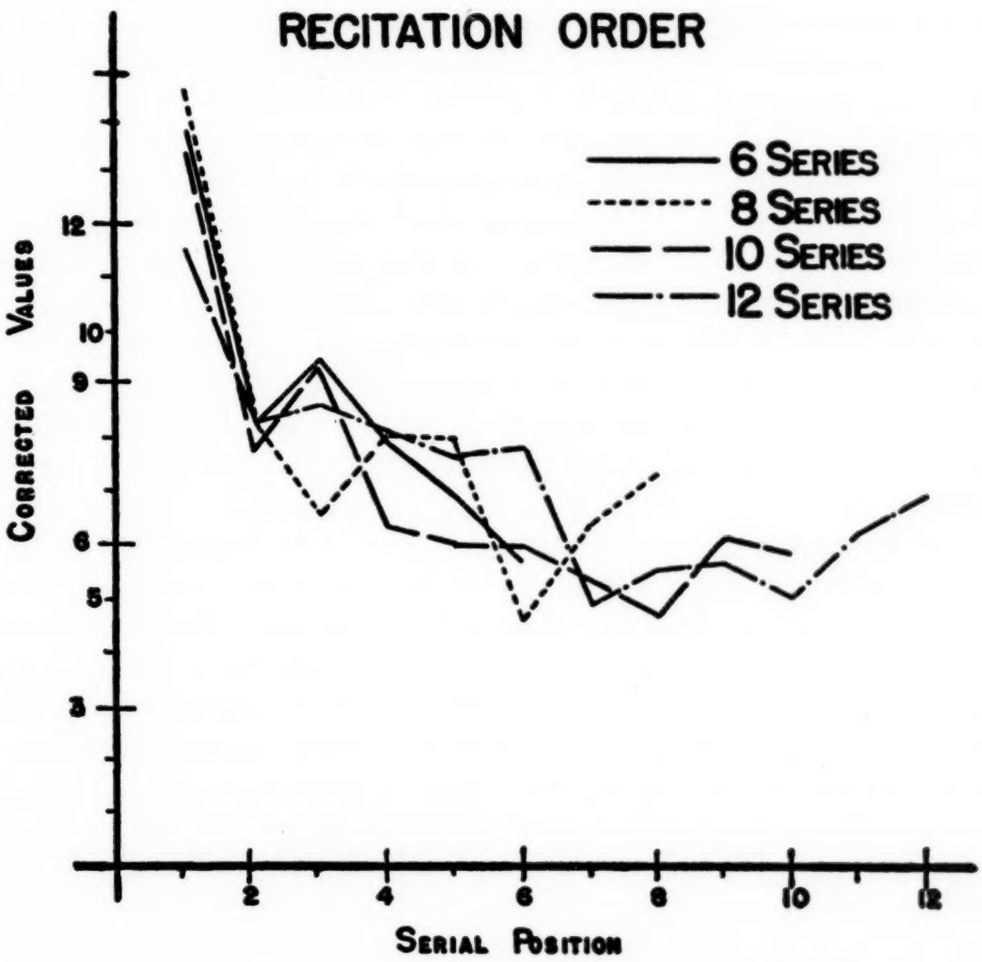

FIG. 2. Variation of difficulty of reproduction according to serial position for the different series. Difficulty measured by the average correct response, per subject per trial per list, after the syllables are weighted for differences in association value. The recitation order of the syllables varies from trial to trial, and from series to series.

Considering the curve of recitation order, however, it is seen that the simple concept of finality is not adequate to account for the findings. Due to rotation in the recitation series, each syllable of the learning series appears in every position of the recitation order. Yet in the recitation curve degree of recall closely follows the recitation order, the curve of correct responses being highest for the initial position and decreasing as we proceed toward the last position. Primacy in the recitation order, then, is clearly the condition most favorable to correctness of recall.

Analyzing this primacy factor we note the following:

1. The average temporal distance between the learning positions of the syllables and their occurrence in the first position in the recitation order is less than the temporal distance to the second position and this distance increases as we move to the last position of the recitation order where it is at a maximum. The factor of relative temporal distance between learning and recitation is then certainly a concomitant if not a causal factor in the recall.

2. An increase in temporal distance brings with it an increase of material interpolated between learning and recall as we proceed from the initial to the last position of the recitation order. Thus the regressive inhibiting influences progressively increase.

3. Finally there must be considered the possibility of proactive effects, since the recitation curve is in appearance typical of those for proaction. We have been considering the inhibiting influences as being interpolated between the learning and the recitation, that is, retroactive inhibition. And to a large extent this is correct. However if we consider the method used in this experiment and use the letter L for learning and R for recitation we find the procedure to be as follows: L_1—R_1—L_2—R_2—L_3—R_3—L_4—R_4 etc. Let us concentrate for the moment on the trial L_3—R_3. The inhibiting influences operating intraserially are mainly of the retroactive type and become less as we proceed toward the final syllable in L_3. But if we consider L_3—R_3 in its position in the total series of trials then we must realize the possible proactive effect upon R_3 of the previous occurrence of L_1—R_1—L_2—R_2. If L_1, and L_2 and L_3 each contained entirely different syllables, we would on the basis of proaction alone expect the learning in R_3 to be inferior to that of R_2 and R_1. Since, however, L_1, L_2 and L_3 are the same, the mere repetition of the syllables operates to lessen this proactive effect so far as a comparison of R_3 with R_2 and R_1 is concerned. Yet within the series R_3, it may be that these proactive influences are still operating. Certainly the shape of the recitation curve is consistent with this hypothesis. Again the said shape is consistent with the idea of an inhibiting influence on succeeding members of the recitation series by the process of recall of each preceding member.

While it is obviously impossible to isolate in this experiment the several causal factors it must nevertheless be held in mind that the findings herein are not inconsistent with the results to be expected from the operation of temporal distance, regressive inhibition and progressive inhibition.

This explanation, developed from the concepts of finality in the learning series and primacy in the recitation series, differs radically from the usual primacy-finality explanation based on one curve.

Although recency in temporal terms is involved there is also connoted the presence of inhibitory forces.

Subsequent to the present investigation Melton and Lackum (9) have published an article proposing a two-factor theory of retroactive inhibition which is in accord with the results of this experiment. Their experiment was performed using the method of interpolated material. The two factors were obtained by the effect of interpolated material on the recall of the original material (retroactive inhibition), and the effect of original material on the recall of interpolated material (proactive inhibition). The authors discuss the serial position factor in their results, but from the point of view of inter-list intrusion, which is not comparable to the serial position factor as discussed in the above experiment. The authors discuss also a two-factor theory "which specifies that a portion of the retroactive inhibition . . . is attributable to something that happens to the original activity during the learning of the interpolated activity (Factor *X*), and the remainder . . . to the competition of the original and interpolated responses at the time of attempted recall and relearning. . . . The total inhibition of the recall and relearning of the second activity (proactive inhibition) would, therefore, be less than the total inhibition of the first activity (retroactive inhibition) (p. 171)." These results of Melton are in accord with those of this experiment, with the superficial differences that the difference in experimental technique would make one expect. The curve of the learning order in Fig. 1 substantiates the retroactive inhibition curve, and the curve of the recitation order in Fig. 2 substantiates the proactive inhibition curve described by Melton. The Factor *X* indicated by Melton is also present in that the values for retroactive inhibition are higher, at least at the terminal points, than the values for proactive inhibition in this experiment. Therefore, we would say that the two factor theory of inhibition as described by Melton will also explain the results of this experiment.

(Manuscript received December 8, 1941)

BIBLIOGRAPHY

1. BRITT, S. H. Retroactive inhibition: a review of the literature. *Psychol. Bull.*, 1935, **32**, 381–440.
2. CRAFTS, L. W. Primacy and recency in the learning of visual diagrams. *Amer. J. Psychol.*, 1932, **44**, 763–767.
3. FOUCAULT, M. Les inhibitions internes de fixation. *L'année Psychol.*, 1928, **29**, 92–112.
4. GIBSON, J. J., & RAFFEL, G. A technique for investigating retroactive and other inhibitory effects in immediate memory. *J. gen. Psychol.*, 1936, **15**, 107–116.
5. GLAZE, J. A. The associative value of nonsense syllables. *J. genet. Psychol.*, 1928, **35**, 255–269.
6. MCGEOCH, J. A., & MCGEOCH, G. Studies in retroactive inhibition VI: The influence of the relative serial position of interpolated synonyms. *J. exp. Psychol.*, 1936, **19**, 1–23.

7. McGeoch, J. A., & McKinney, P. Studies in retroactive inhibition VIII: The influence of the relative order of presentation of original and interpolated paired associates. *J. exp. Psychol.*, 1937, **20**, 60–63.
8. McGeoch, J. A., & Sisson, E. D. Studies in retroactive inhibition XI: The influence of the relative serial positions of interpolated synonyms in twenty-item lists. *J. exp. Psychol.*, 1938, **22**, 547–554.
9. Melton, A. W., & von Lackum, W. J. Retroactive and proactive inhibition in retention: Evidence for a two-factor theory of retroactive inhibition. *Amer. J. Psychol.*, 1941, **54**, 157–173.
10. Mitchell, D. Errors in the memorization of numbers. *Amer. J. Psychol.*, 1935, **48**, 1–16.
11. Mitchell, D. Anticipatory place-skipping tendencies in the memorization of numbers. *Amer. J. Psychol.*, 1934, **46**, 30–39.
12. Philip, B. R. Proactive and retroactive effects in the recognition of form. *J. exp. Psychol.*, 1940, **26**, 502–513.
13. Robinson, E. S., & Brown, M. A. Effect of serial position upon memorization. *Amer. J. Psychol.*, 1926, **37**, 538–552.
14. Sand, M. C. The effect of length of list upon retroactive inhibition when degree of learning is controlled. *Arch. Psychol.*, 1939, **33**, No. 238.
15. Schmeidler, G. R. Retroaction and proaction in serial learning. *Amer. J. Psychol.*, 1939, **52**, 592–601.
16. Welch, G. B., & Burnett, C. J. Is primacy a factor in association-formation? *Amer. J. Psychol.*, 1924, **35**, 396–401.

THE STAPEDIUS MUSCLE IN RELATION TO SOUND CONDUCTION [1]

BY ERNEST GLEN WEVER AND CHARLES W. BRAY

Princeton University

In a previous experiment (3) the function of the tensor tympani muscle was studied by applying artificial tension to its tendon and observing, by means of the electrical responses of the cochlea, the resulting changes in sound transmission. In the present investigation similar methods were employed for the stapedius muscle.

The first attempt to study the stapedius muscle by imitating its action was that of Mach and Kessel in 1872 (2). They worked on cadavers, and observed through a microscope the movements of the head of the malleus and the round window membrane during stimulation with sound. After wide exposure of the middle ear, a thread was looped around the stapes and then led over a pulley to a weight. It was found that the application of tension reduced the amplitude of vibratory movements.

In 1935 Hughson and Witting (1) reported experiments in which the electrical responses of the ear were used to measure the effects of tension on the stapedius tendon. In anesthetized cats a small hook was attached to the stapedius tendon, and tensions were applied during sound stimulation. The result of the tension was a marked reduction in transmission.

Method and Procedure

In the present experiments, as in those previously reported on the tensor tympani muscle, cats were used as experimental animals. They were first anesthetized with Dial, and then were curarized to the point at which respiration ceased and reflex movements of the pinna in response to pinching could no longer be elicited. Previous tests have shown that reflex action of the tympanic muscles is absent for this degree of curarization. The animals were maintained under artificial respiration.

The stapedius tendon and muscle were exposed by careful dissection in the mastoid region. The bone was removed with fine rongeurs and by grinding with a dental burr. The excavation was carried to the fissure in which the stapedius muscle lies, and extended laterally at the level of the muscle. A fine thread was prepared by twisting a few silk fibers and waxing, and this thread was passed around the stapedius tendon, the end tied in a slip knot, and the knot carefully secured about the tendon. The thread was carried over a pulley to a scale pan. The direction of pull was adjusted so as to reproduce as nearly as possible that exerted normally by the muscle; this direction varied slightly in different experiments, but was estimated as within 10 degrees of the normal.

[1] This investigation was aided by a grant from the Rockefeller Foundation.

In one experiment this technique was varied, and the thread was tied around the neck of the stapes at the level of insertion of the tendon.

The animal was stimulated with pure tones from 100 to 10,000~, at various intensities. The electrical responses of the cochlea were picked up with a silver foil electrode on the round window membrane, and after amplification were measured with a General Radio Type 736-A wave analyzer. Numerous tensions were used up to that which broke the tendon or pulled the stapes away from its attachments. The experiments were carried out on 6 cats.

Results

The usual result of the application of tension to the stapedius tendon is a reduction of transmission, as illustrated in Fig. 1. In

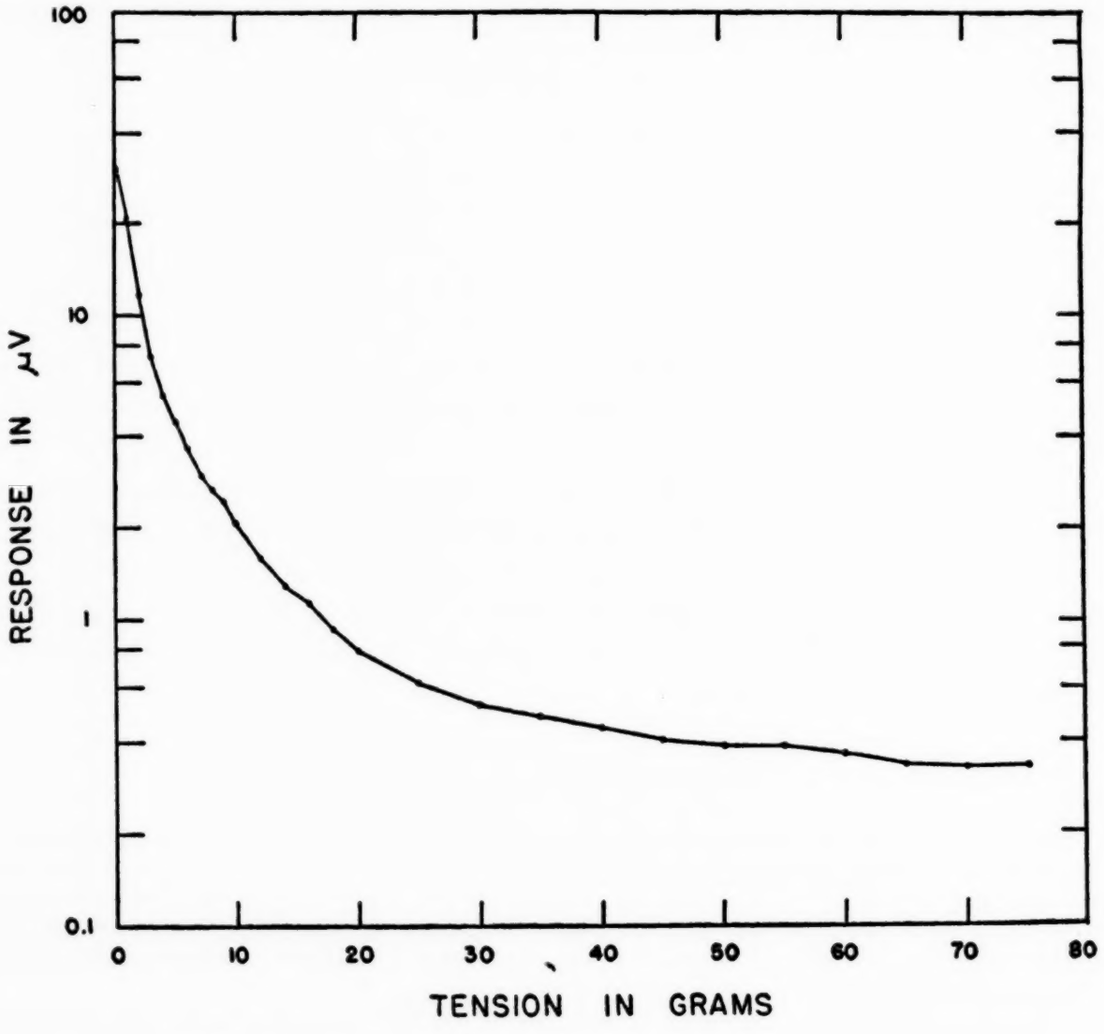

Fig. 1. Effect of tension on the stapedius tendon upon responses to a 1000~ tone. Cat No. 437 L.

this experiment the stimulating tone was 1000~, presented at an intensity which in the absence of tension yielded a response of 30 microvolts. The effects of tensions up to 75 grams are shown. As the load is varied from zero upwards the responses fall rapidly at first, and then progressively more slowly. At 40 grams the reduc-

tion of response amounted to 36.5 db, but further increase of tension had comparatively little effect. The greatest reduction was 39 db, obtained for a tension of 75 grams. This tension was the maximum that could be applied, as immediately after the reading was obtained the tendon was pulled away from its attachment to the stapes.

The breaking strength of the tendon varied considerably among the animals. In four other animals the tendon was pulled loose at tensions of 41, 60, 81, and 100 grams, respectively. In three of these the responses returned to normal afterwards. In the remaining animal, in which the tendon broke at 81 grams, the responses fell rapidly to 20 db below the normal value, showing damage to the ear. In the one animal in which the thread was tied around the stapes the maximum tension was 81 grams, which pulled the stapes out of the ear and caused a serious loss of responses.

The effect of the application of tension to the stapedius tendon upon the intensity function of the cochlear response is shown in Fig. 2. Each curve represents, for a given constant tension, the relation between cochlear potential and stimulus intensity. The tension in grams is indicated by the numbers near the lower ends of the curves.

The principal effect of tension is merely to displace the functions to the right along the intensity axis. Below the region of overloading the curves remain straight, and are of the same slope. In the region of overloading, however, slight variations appear. For the higher tensions the curves bend somewhat more rapidly, and approach a maximum at a lower level. This observation suggests an effect of the tension beyond a simple reduction of transmission.

The effects of tension upon responses to various tones are shown in Fig. 3. Here 9 curves are drawn, for tones from 100 to 5000~, for each of which the stimulus intensity was adjusted to a value that for zero tension yielded a response of 30 μv.

For most of the tones the effects are like those already indicated in Fig. 1. For three of them, 1500, 1750, and 2000~, however, the smaller tensions give a preliminary improvement of response. As the tension is varied from zero upwards the curves rise sharply to a maximum and then suffer a rapid decline.

Marked differences are shown elsewhere in the frequency range. Generally speaking, the low tones are affected most, and the high tones least. Thus a tension of 40 grams caused a reduction of 46 db for 200~, and only 17 db for 5000~. However, the relation is not strictly in order of frequency, and it varies somewhat for different degrees of tension.

The relations to frequency are further revealed in equal-response curves like those of Fig. 4. Here each curve indicates, for a given condition of tension, the stimulus intensity required at each of 13

frequencies to give a standard response of 10 μv. The larger effect of tension on the low tones is shown by the greater separation of the curves. This effect is especially noticeable for a tension of 3 grams,

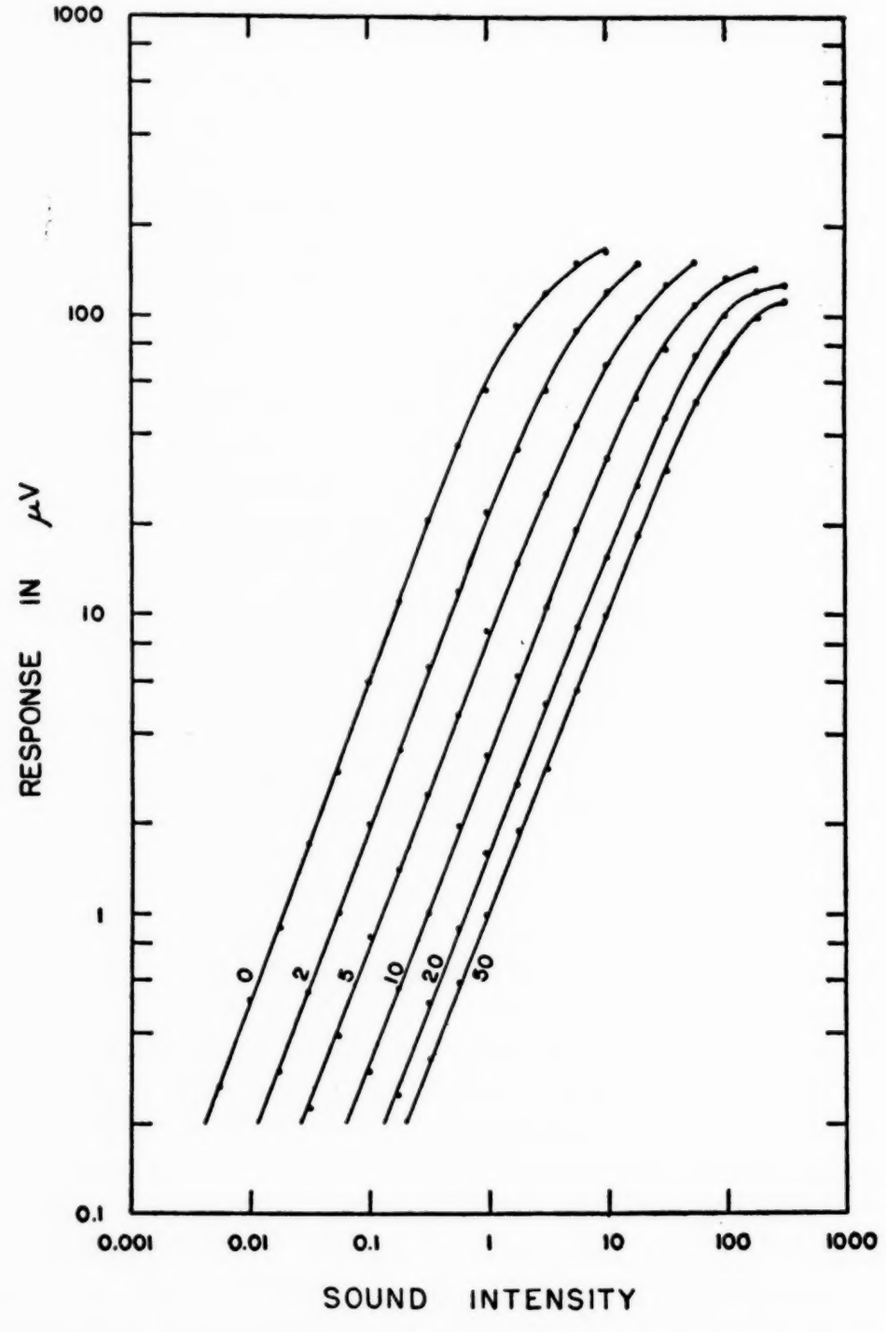

FIG. 2. Intensity functions for a 1000~ tone in the presence of tension on the stapedius tendon. The abscissa shows sound intensity in dynes per sq. cm. The numbers near the lower ends of the curves indicate the tension in grams. Cat No. 434 L.

which has marked effects for tones of 1500~ and below, and only slight effects for higher frequencies. At 2000~ there is a slight improvement.

Relation to Distortion.—It has often been suggested that contraction of the stapedius muscle should alter the symmetry of the ossicular system and thus affect the distortion of the ear. Tests were

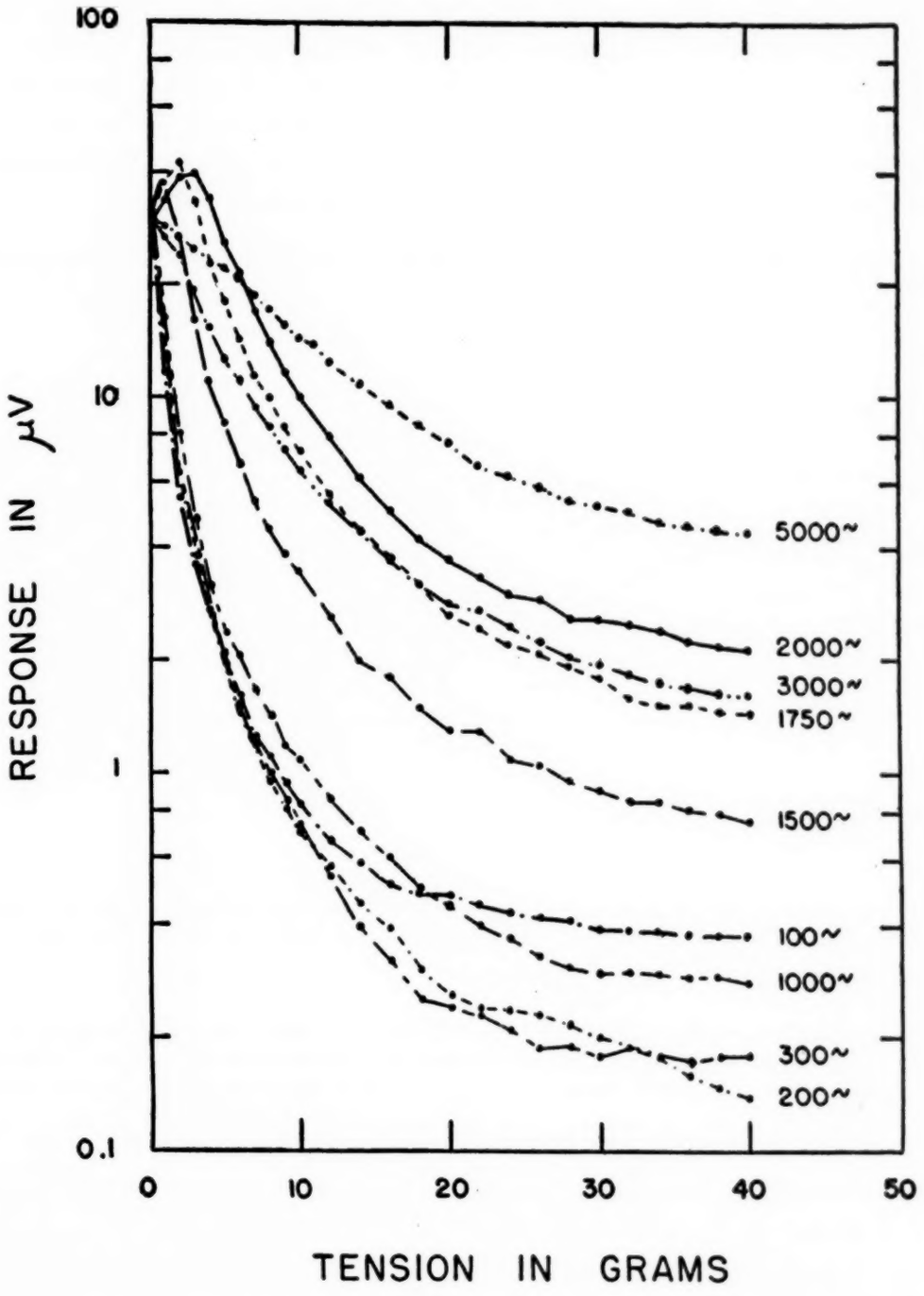

FIG. 3. Tension curves for 9 stimulus frequencies. Cat No. 437 L.

carried out to determine the effects of artificial tension upon the distortion pattern. Some of the results are given in Fig. 5. The stimulus was a pure tone of 1000~, and for each condition of tension

the intensity was adjusted so as to produce a response in which the fundamental component was 70 μv., as indicated by the uppermost curve of the figure. Other components in the response up to the 5th harmonic were measured, and their magnitudes are shown by the curves below.

The curves exhibit large variations as a function of tension. As the tension is varied from zero upwards the harmonics are at first reduced, and then rapidly augmented. However, it is evident that the different components do not follow exactly the same courses; at times one is reduced while another is increased.

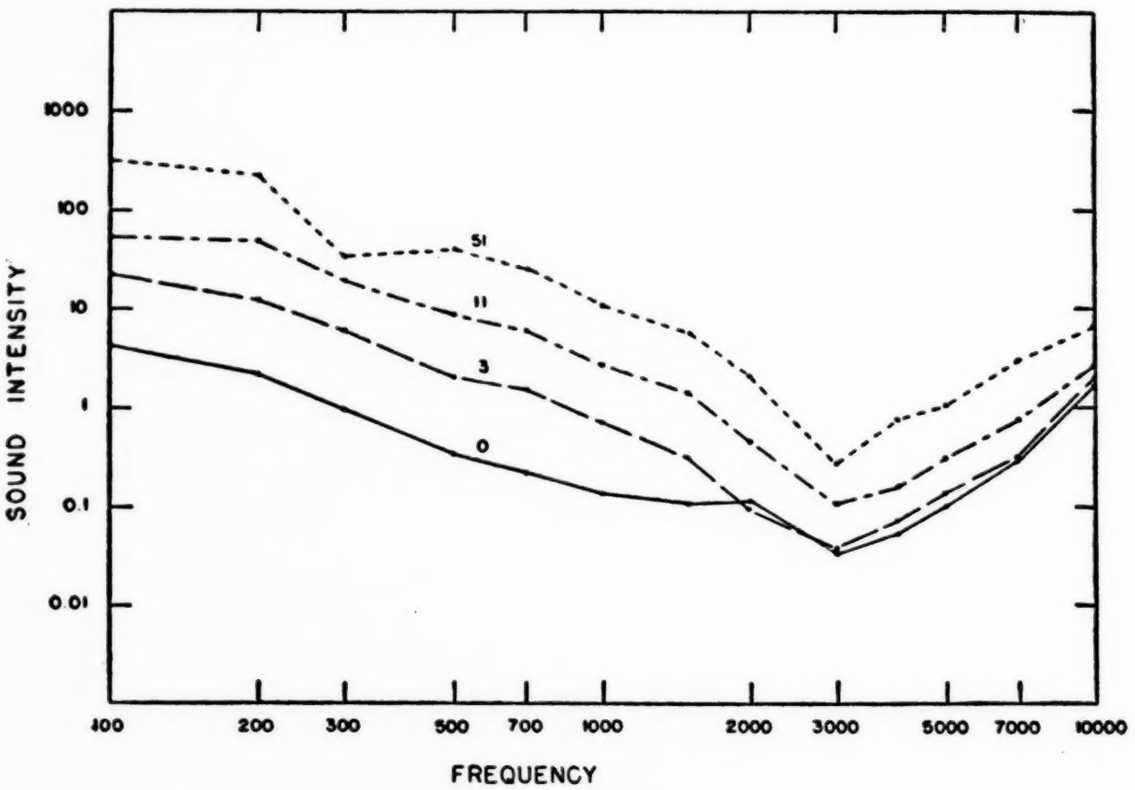

FIG. 4. Effect of tension upon sensitivity. Each curve shows the sound intensity in dynes per sq. cm. required at various frequencies to produce a standard response of 10 μv. The tension on the stapedius tendon, which is constant for each curve, is indicated by the number above the curve. Cat No. 432 L.

Measurements on other animals likewise showed changes in harmonic content as a result of tension on the stapedius tendon, but the form of the changes varied. In another series of tests, for example, the smaller tensions gave usually an irregular rise in the harmonics, and the larger tensions produced no further change.

The effects do not seem to be related to response level, provided that the region of severe overloading is not reached. When overloading is great, however, the application of tension to the stapedius tendon has relatively little effect upon the harmonic pattern.

Observations made on one animal in which the thread was tied around the neck of the stapes gave results essentially similar to those described above.

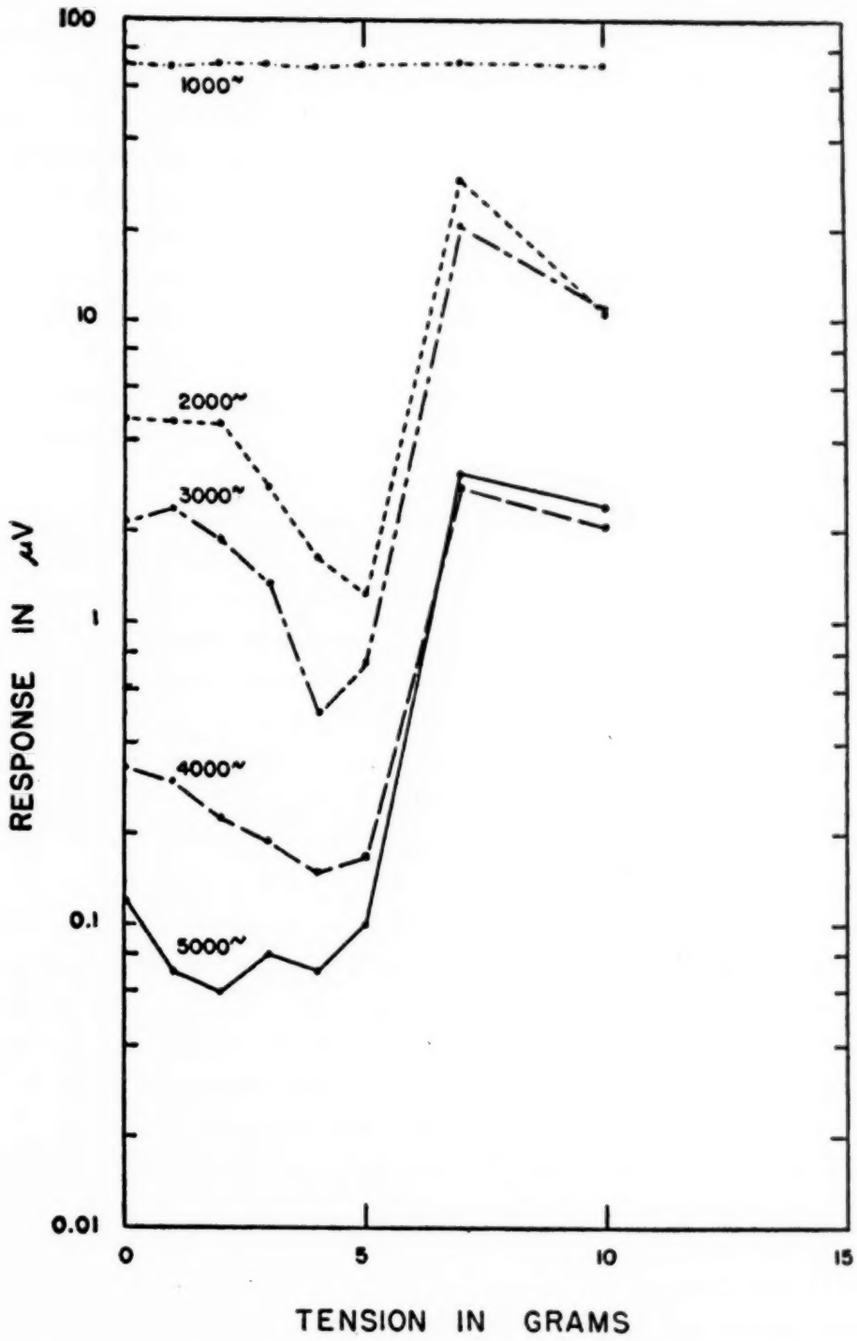

FIG. 5. Effect of tension upon distortion. The results represent a wave analysis of the response obtained to a pure tone of 1000~. The fundamental component is indicated by the uppermost curve, and the first four overtones by the remaining curves. Cat No. 437 L.

DISCUSSION

The results of this experiment are in general support of the hypothesis that the stapedius muscle, like the tensor tympani, is protective in function, and reduces the transmission of vibrations to

the inner ear. An exception to this generalization appears for certain tones of the middle range, which are improved in transmission by the exertion of slight tension. As tension is increased beyond about 4 grams, however, the conduction of these tones also is reduced.

A comparison of the present results with those obtained similarly for the tensor tympani muscle shows significant differences. A given load applied to the stapedius tendon is about 10 times as effective as when applied to the tensor tympani tendon. For example, in results previously reported for the tensor tympani, a load of 20 grams reduced responses to a 1000~ tone from 30 to 7 μv., a change of about 4-fold or 12 db, whereas Fig. 1 shows for the same load on the stapedius tendon a reduction from 30 to 0.8 μv., a change of 37-fold or 31 db. It is evident that despite its smaller size and strength the stapedius muscle takes a significant share of the function of protecting the ear against overstimulation. Indeed, it may be the more important of the two muscles in this respect.

The relations to frequency are somewhat different for the two tympanic muscles. For the tensor tympani muscle the application of tension reduces transmission and shifts the sensitivity curve upward along the frequency scale. It has been suggested that tension adds mechanical resistance and stiffness to the ossicular system, thereby both damping its movement and raising its natural frequency.

For the stapedius muscle there is no noticeable shift of the region of maximal sensitivity, which under tension remains about 3000~. There is, however, a significantly greater impairment of tones below this region than of tones above it, and within the region itself there is a slight attunement. It is reasonable to suppose that tension on the stapedius tendon, as on the tensor tympani, adds both damping and stiffness to the middle ear apparatus. From the magnitude of the reduction of transmission in comparison with that shown for the tensor tympani it appears that a relatively large amount of the force is effective in damping the ossicular movements.

The effects of tension on the stapedius tendon upon distortion are of interest in relation to the general problem of the operation of the middle ear mechanism. Other experiments (**4**) have shown that under normal conditions this mechanism operates with high fidelity in the transmission of sound. Distortion is a prominent feature of the response of the ear, especially when the stimuli are strong, but its chief locus is the inner ear. No significant changes in the distortion pattern have been observed as a result of tension on the tensor tympani tendon, or of alteration of air pressure in the middle ear cavity (**5**). Tension on the stapedius tendon is the only procedure so far found by which distortion may be appreciably altered.

That distortion may be reduced as well as increased by tension on the stapedius tendon calls for special consideration. Under some

conditions the total harmonic content of the response may be reduced to one-third the normal amount, or increased more than seven-fold. Particularly striking is the rapid change in distortion with alteration in the degree of tension; in the results of Fig. 5, for example, an increase from 5 to 7 grams alters the harmonic content from well below normal to the maximum value shown. These relations suggest that the effect of tension on the stapedius tendon is to introduce distortion into the middle ear mechanism, and that this distortion may summate with or partially cancel the distortion which arises in further processes. A variation in the phase relations of harmonics produced in separate distortion processes can readily account for the alterations of total content, and in the rapid changes above and below the normal level.

It is difficult to assign any definite reason that tension on the stapedius tendon introduces distortion, while tension on the tensor tympani tendon and change of middle ear pressure have no discernible effect. We may of course point out the anatomical peculiarities of the stapedius attachment; tension on its tendon rocks the stapes outward on a pivot formed by the posterior portion of the annular ligament at the oval window. Such a change may well alter the symmetry of the mechanical system in such a way as to introduce distortion.

Summary

The function of the stapedius muscle was investigated by artificially reproducing its action and observing the effects upon the electrical responses of the cochlea. Tension applied through a thread attached to the stapedius tendon causes a marked reduction in transmission for low tones, a smaller reduction for high tones, and for certain tones of the middle range a preliminary improvement and then a considerable decline. The results thus show that in general the stapedius muscle, like the tensor tympani, operates so as to protect the inner ear against overstimulation.

A further effect of tension on the stapedius tendon is a significant alteration of the pattern of distortion of the ear.

(Manuscript received December 11, 1941)

References

1. Hughson, W., & Witting, E. G. An objective study of auditory fatigue. *Acta Oto-laryngol.*, 1935, **21**, 457–486.
2. Mach, E., & Kessel, J. Versuche über die Accommodation des Ohres. *Sitzungsber. Akad. Wiss., Wien, Math-nat. Cl.*, Abt. 3, 1872, **66**, 337–343.
3. Wever, E. G., & Bray, C. W. The tensor tympani muscle and its relation to sound conduction. *Ann. Otol., etc., St. Louis*, 1937, **46**, 947–961.
4. Wever, E. G., Bray, C. W., & Lawrence, M. The locus of distortion in the ear. *J. acoust. Soc. Amer.*, 1940, **11**, 427–433.
5. Wever, E. G., Bray, C. W., & Lawrence, M. The effect of middle ear pressure upon distortion. *J. acoust. Soc. Amer.*, 1941, **13**, 182–187.

STUDIES IN SERIAL VERBAL DISCRIMINATION LEARNING. I. REMINISCENCE WITH TWO SPEEDS OF PAIR PRESENTATION [1]

BY DAVID C. McCLELLAND

Connecticut College

INTRODUCTION

The phenomenon of reminiscence [2] has been the subject of careful and extensive investigation by Ward (17) and Hovland (4, 5, 6, 7). It is defined in these studies as an improvement in the recall or relearning of a partially learned task after periods of rest in which no practice could have occurred. It is the antithesis of forgetting, which is poorer recall or relearning following periods of no practice.

The hypothesis which has been consistently favored by these investigators as an explanation for reminiscence is that rest periods "provide an opportunity for the dropping out or forgetting of wrong associations . . . less strongly formed than the right ones and hence forgotten more rapidly" (14, p. 322). In serial learning such an explanation (called the differential rate of forgetting hypothesis) readily applies. The remote associations formed in learning nonsense syllables by rote may be considered the weak wrong associations which drop out more rapidly with time than the immediate first order associations. The net result of this differential rate of forgetting would be an initial improvement in the recall of the correct first order associations after a rest period—an improvement which satisfies the requirements of the definition of reminiscence. There is even some evidence that reminiscence appears *only* in situations in which the wrong response tendencies are remote associations. Hovland (7) failed to find reminiscence in paired-associate learning presumably because remote associations could not be formed that might dissipate rapidly in a rest interval.

This explanation of reminiscence depends upon an assumed dif-

[1] This paper is a modification of a thesis presented to the University of Missouri in partial fulfillment of the requirements for the M.A. degree. I am indebted to Dr. A. W. Melton for his invaluable assistance in the planning and execution of the research, to Dr. Carl I. Hovland for his aid in the interpretation of the results, and to Dr. John P. Seward for his helpful suggestions in the final preparation of the report.

[2] The reminiscence investigated by Ballard (1) and G. O. McGeoch (13) *et al.* will be considered for the present a separate phenomenon and will not be taken into consideration in the theoretical discussions throughout this series.

ferential in the rate of forgetting of right and wrong responses. Such an assumption appears to have some basis in fact if wrong responses are always weaker. For the work of Krueger (**11**), A. C. Youtz (**20**) and others indicates that weaker learned responses are forgotten more rapidly. Direct confirmation of the differential rate, however, is lacking. So the present series of experiments was undertaken in an effort to test separately the rates of forgetting of right and wrong parts of a learning situation.

Since serial learning did not lend itself easily to such a test, a form of discrimination learning was adopted in which there are two sets of responses, the retention of which can be measured separately. In this technique words are presented in pairs at regular intervals. The subject has to learn to pick the right word in each pair. A bell rings if he chooses the right word. The subject's task, after he makes a random selection on the first trial, is to remember which word in each pair rings the bell. He has to learn to say one word (the one followed by the bell) and not to say the other in each pair.

The first experiment in the series using this technique was undertaken to discover if reminiscence could be obtained under the usual conditions of learning by rote if discrimination instead of serial association was the task involved. Because of an important difference in the nature of the wrong responses involved, reminiscence would not be expected in discrimination learning according to the differential rate of forgetting hypothesis.

The wrong responses in serial learning are positive interfering tendencies (remote associations) *acquired* in the course of learning; in discrimination learning the wrong response is saying the wrong word—a response which exists on the first trial at its maximum strength and which *declines* in strength as it is replaced either by an acquired avoidance response to the wrong word or an acquired 'approach' response to the right word. Forgetting of either of these responses acquired in the course of learning will reinstate in strength the wrong response—an increase in strength which will result in net forgetting and not the net improvement required to explain reminiscence. So far as can be ascertained there is no acquired wrong response in discrimination learning which, by being forgotten rapidly, could lead to reminiscence. Consequently if reminiscence is obtained in discrimination learning, it cannot be caused by the same differential rate of forgetting that purports to explain reminiscence in serial learning.

A secondary aim of the first experiment was to test Hovland's hypothesis that reminiscence is only to be found when there are remote associations to be forgotten rapidly with rest. He attributed his failure to obtain reminiscence in paired-associate learning

(7) to the fact that a counterbalanced order of presentation prevented the formation of such remote associations. In the present experiment the pairs are counterbalanced in similar fashion. If reminiscence appears under these conditions, remote associations cannot be considered essential for its appearance.

Experimental Procedure

The essential apparatus was a modified Hull memory-drum for presentation of the pairs of adjectives and a voice-key set-up for measuring reaction times. Two lists were made of 20 pairs of 2-syllable dissimilar adjectives. These were typed on a long circular band of paper in six different random orders. The pairs were spaced so that when the continuous band was hung on the drum, each turn of the drum would bring one pair into the window. A polygraph recorded the reaction times on a moving sheet of waxed paper in a room near the experimental one. It gave readings to .001 second accurately.

The subjects were instructed to choose one word of each pair and say it out loud beginning with the very first pair that came into the window. If their choice happened to be right, the experimenter rang a bell. If not, it was wrong and nothing happened. On the second trial they were to try and remember which words of each pair had rung or would have rung the bell on the preceding trial and to choose those again. They were told it was a memory experiment only and that there was no rational system by which they could figure out which words were going to be right. The criterion of mastery was two successive trials in which they selected the correct adjective from each pair. Since the subjects found this very difficult to attain with the shorter presentation interval, mastery was reduced for these subjects to one perfect recitation.

The bell was placed in a circuit open in two places so that it would ring when a key was pressed by the experimenter and when a contact was made by a cam revolving on the drive shaft of the drum motor. This contact was so placed in the revolution of the cam that if the experimenter pressed his key, the bell would ring at a fixed time after the end of the exposure interval with constant duration and intensity. The end of the exposure interval was marked by the fall of a shutter also controlled by a cam. It remained down covering the pair while the rewarding bell was rung. For groups A and B there was an exposure interval of 2.3 seconds and a reward interval (when the shutter was down) of 2.3 seconds. These figures include the time the drum took in turning. So for all practical purposes their sum may be referred to as a 4-second presentation interval. For groups C and D these values were 1.65 and .65 seconds respectively. Their sum may be referred to as the 2-second presentation interval.

The system of correct responses was randomized so that the correct choice came from the top or bottom of a pair in an irregular order. It also varied from trial to trial since the words in a pair were alternated from top to bottom, and since each pair appeared in a different serial position in each of the six orders. To further counterbalance any consistent effects of a reward system, half of the subjects were run with opposite members of the pairs correct.

Each subject served two days, the first for practice on which he learned one list through without interruption to mastery. On the second or work day the 96 subjects were divided into 4 groups. The reminiscence groups (B and D, 24 subjects each) learned to a criterion of 15 out of 20 correct responses. Then they were interrupted for 2 minutes. During this time they were asked to step down and take a test of steadiness as a measure of muscular tension. At the end of the rest period they resumed the learning and continued it until they reached mastery. The control groups (A and C, 24 subjects each), when they reached the criterion of 15 correct choices in 20, continued without interruption on to mastery.

Groups A and B learned at a presentation rate of 4 seconds including reward time. There were 3 blank spaces or a total of 12 seconds between trials. For Groups C and D the conditions were identical except that learning proceeded at a 2-second rate of presentation. This reduced the rest between trials to 6 seconds.

The subjects used were 96 naive students of both sexes in elementary psychology at the University of Missouri who were required to serve in at least one experiment as the laboratory equivalent of a term paper.

Results

4-second Rate, Groups A and B.—The data are summarized in Table 1. There is no difference between the control and 2-minute rest groups before or after the rest. The introduction of the pause has no noticeable effect either on recall or relearning scores that cannot be explained by the normal increment from one trial to the next. The gain in the control group is if anything slightly greater indicating that rest produced some slight forgetting. Reaction time data yielded similar results. The mean choice reaction times [3] on OL_2 were 1.070 sec. and 1.017 sec. for the control and rest groups respectively. The mean reaction times on the first trial of relearning were .990 sec. and .976 sec. respectively. The gain (shown as a decrease in reaction time) in the control group was .08 sec., a figure

TABLE 1

Mean Scores in Original Learning and Relearning Learning at 4-second Rate of Presentation

Group	N	OL_1	OL_2	Recall	Relearning
A. Control	24	5.29	15.92	16.67	5.92
σ_m		.36	.16	.30	.52
B. 2-min. rest	24	5.67	15.63	16.29	6.75
σ_m		.39	.18	.38	.54

OL_1 = Trials to reach the criterion of 15 in 20 correct choices.
OL_2 = Number of correct choices on the criterial trial.
Recall = Number of correct choices on the first trial after the criterial trial.
Relearning = Number of trials to reach mastery of 2 perfect errorless trials in succession (counting from OL_1).

almost exactly like the rest group gain of .041 sec. Apparently rest had no differential effect on choice reaction time either.

2-second Rate, Groups C and D.—Table 2 and Fig. 1 summarize the results obtained when the presentation rate was increased to one pair every 2 seconds. The measures are the same as in Table 1 except that relearning is to a criterion of one perfect trial instead of two. The 2-minute rest pause had no significant effect on recall but shortened relearning by 3.62 trials. Data on the mean number of trials to reach lesser criteria such as 18 and 19 out of a possible 20 correct choices are shown graphically in Fig. 1. In original learning the learning curves of the two groups of 24 subjects are very close together but after the rest is introduced they begin to draw apart. As the criterion becomes stricter the advantage of the rested group

[3] The mean reaction times represent 480 determinations each.

TABLE 2

MEAN SCORES IN ORIGINAL LEARNING AND RELEARNING LEARNING AT A 4-SECOND RATE OF PRESENTATION

Group	N	OL_1	OL_2	Recall	Relearning
C. Control	24	8.46	15.75	14.38	13.75
σ_m		.85	.13	.45	1.24
D. 2-min. rest	24	8.46	15.63	14.92	10.13
σ_m		.72	.18	.53	1.13
Difference				.54	3.62
$\sigma_{diff.}$				.70	1.68
Critical ratio				.77	2.16

Symbols the same as in Table 1 except that relearning is to a criterion of 1 perfect trial.

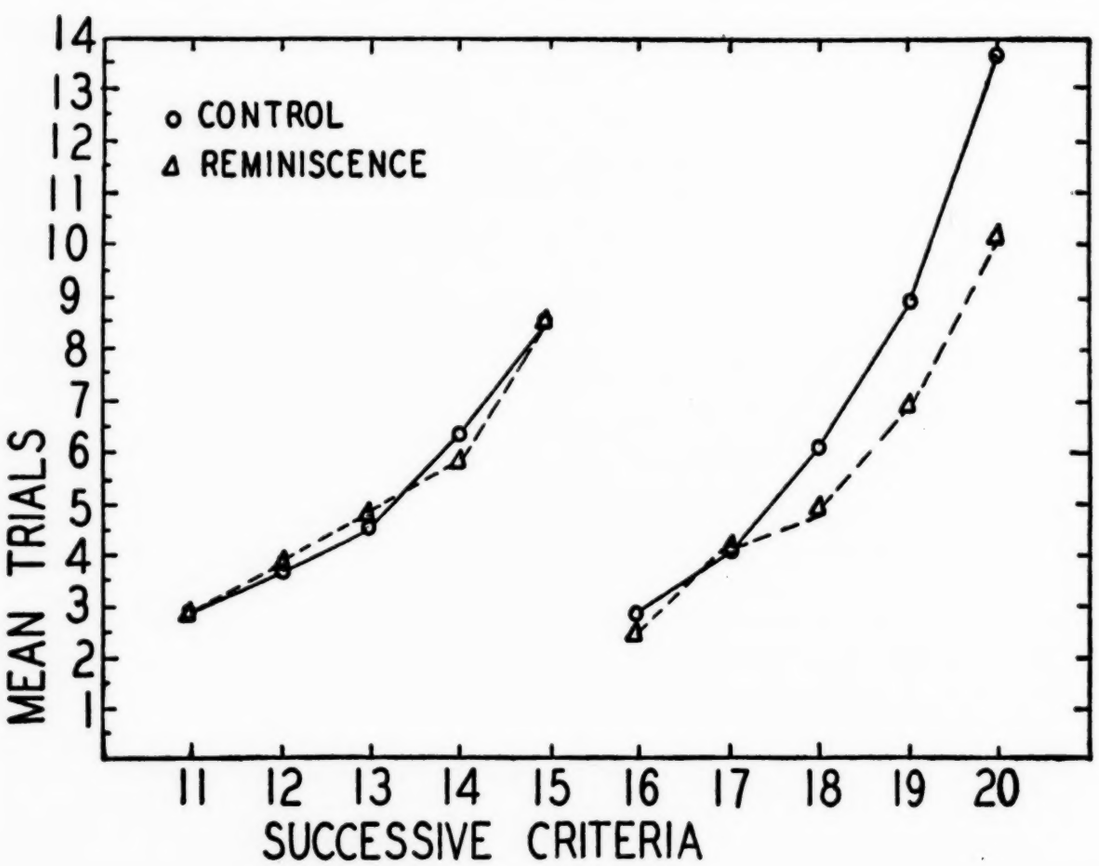

FIG. 1. Groups C and D, 2-second rate of presentation. Average trials to reach successive numbers of correct responses out of 20 before and after the original learning criterion of 15 out of 20.

becomes greater and greater until the difference reaches the point of significance at one perfect trial ($P < .05$).[4] Extrapolation of this trend would suggest that the stricter criterion of 2 perfect trials in

[4] This is to be read: A difference this large in either direction could have occurred by chance less than 5 out of 100 times.

succession, had it been required, would have produced an even greater advantage. The reaction time data, which were computed only for recall scores, and which are not reproduced here, showed no significant trends.

A comparison of trials taken to reach the criterion of one perfect trial in the control groups (A and C) shows clearly the efficacy of the longer presentation interval. It took control group A learning at a 2-second rate an average of 9.42 more trials to make one perfect recitation than it did group C learning at a 4-second rate to the same criterion.[5] This difference is highly reliable with a critical ratio of 7.25 ($P < .01$). The results of Hovland's experiment (**5**) with varying speeds of syllable presentation in serial learning are confirmed and extended to discrimination learning.

Discussion

The method of experimentation used in this study is related to Peterson's mental maze (**15**) and particularly to a modification of it used by Van Tilborg (**16**). It has been standardized somewhat more than either of these by controlling the presentation rate and the nature of the reward. Such a technique has the important advantage that it is well-adapted to studies of recognition memory and discrimination because its procedure is like the standard procedure in learning by rote. That this similarity is not merely formal but functional is indicated by the fact that increasing the rate of presentation from 4 seconds to 2 seconds produced reminiscence after a 2-minute rest just as it did in Hovland's study with serial learning (**5**). Likewise the slower presentation rate increased the speed with which the material was mastered in both experiments. Since learning appears to proceed under the two techniques according to some of the same laws, it is plausible to expect that results obtained by the discrimination method can be related to the large systematic body of knowledge about serial rote learning (**9**).

The advantage of having two similar methods of measurement before comparing such different mental functions as recognition and reproduction has not always been realized. For example, Luh (**12**) concluded that reproductive memory declined more rapidly than recognition memory despite the fact that the techniques he used to measure each kind of retention differed considerably. The result he obtained ought to be reëxamined using the discrimination method of measuring recognition memory—a method strictly comparable to the rote learning that Luh used to measure reproductive memory.

[5] The mean number of trials taken to reach *one* perfect recitation for groups A and B is not given in Table 1. For group A it was 4.33.

Apart from the advantage to be gained from the fact that it measures a different mental function by a standard controlled-presentation technique, discrimination learning may be recommended as a method which can be conveniently put into use. About 95 percent of the subjects tested learned to select the correct adjective in each of the 20 pairs in about the same number of trials that comparable serial verbal lists take. The technique has the one practical disadvantage that long lists must be used to lower the possibilities of chance attainment of perfect trials. It differs from most learning materials in that about half of the responses are right by chance on the first trial. And thereafter the score on any trial is a function not only of the amount learned but of random successes and failures as well. Consequently chance should contribute to the variability of the scores more than in an experiment in which reproduction rather than choice is required. There is some evidence that standard deviations are larger than in comparable serial verbal lists but not significantly so.

The failure to find trends of importance in the reaction time data suggests that such data may not be of so much value as predicted by Ward (**17**) and others (**9**), especially in view of the labor involved in recording and computing reaction times every few seconds.

The main significance of these results is that reminiscence was obtained in a situation to which the usual explanation of reminiscence in terms of differental rates of forgetting does not apply. This means either (1) that there is actually some wrong response in discrimination learning that has not been identified, or (2) that reminiscence in discrimination learning has a different explanation from that in serial learning, or (3) that the differential rate hypothesis is not the correct explanation of any kind of reminiscence.

(1) The absolute *sine qua non* of a differential rate of forgetting explanation is a wrong response which by dropping out faster than right responses during rest can lead to reminiscence. Furthermore such a response must be *acquired* along with the right response, if it is to be forgotten in the short rest period. This important characteristic of the wrong response has apparently been overlooked in applications of the differential rate hypothesis to situations other than serial learning. In pursuit learning, for example, if the wrong responses result from general awkwardness caused by making the usual angular rather than circular motor coördinations, they cannot be said to be acquired wrong responses. They are rather erroneous habits transferred from normal previous experience. As such they exist at maximum strength in the beginning and decline in strength as they are replaced by correct circular coördinations. That such general habits should be forgotten in a short rest period is not at all

likely especially since normal angular coördinations are actually practiced in the rest periods by most subjects. If these interfering responses are not forgotten but practiced, if anything, during a rest, how can they produce the reminiscence obtained by Humphreys (**10**) and others in pursuit learning?

This does not mean that wrong associations cannot be formed *during* pursuit learning. It does mean that many wrong responses are not so formed and hence cannot be subsequently forgotten as the hypothesis demands. It also means that investigators who use the differential rate of forgetting to explain reminiscence or the efficacy of distributed practice ought to point out the wrong responses that are or may be acquired in the course of learning.

In discrimination learning there do not appear to be any wrong responses that fit this requirement. It is possible that such an association as that between seeing the wrong word and saying the wrong word would suffice. Forgetting of this connection might produce net improvement for the right connection (*i.e.* seeing the right word and saying the right word). Such a wrong association would have to be formed during learning and be strengthened in some way so that it would persist long enough to have sufficient potency for dissipation of it with rest to be beneficial. In discrimination learning strengthening of the tendency to say the wrong response can only be made at the expense of the tendency to say the right response. Therefore if the wrong connection is strengthened at all (that is, if it exists in any amount above chance), it must be stronger than the right connection, if the two responses existed at equal (chance) strength in the beginning. But according to the hypothesis weaker responses are forgotten more rapidly—a circumstance which in this case would mean that the right associations are forgotten more quickly leaving a net gain for the wrong associations after rest instead of reminiscence.[6] No matter how such analyses are made of what constitutes the wrong response in discrimination learning the differential rate of forgetting hypothesis cannot readily be made to predict that reminiscence would be obtained.[7]

[6] It is possible of course that the wrong responses are forgotten more rapidly not because they are weaker but because they are wrong—an eventuality that will be treated in later papers in this series.

[7] The possibilities have not been exhausted, however. Another way in which wrong associations might be formed is by generalization. A given wrong word might be chosen because it is similar to a right word elsewhere in the list. Forgetting of such generalized wrong response tendencies might cause reminiscence. This is not likely because the words in the list were chosen to be as dissimilar as possible both in meaning and form so as to prevent such generalization.

But it is a possibility. The explanation would have to continue that forgetting of the generalized response during rest is caused by forgetting of the primary response. For this process to result in a net gain the generalization gradient would have to be steeper the weaker the primary conditioning—an assumption which has been considered possible by Hull (**8**). That is, the loss in

Further argument against applying the differential rate hypothesis to this situation is to be found in the fact that the one incidental response that is acquired in this situation is a weak avoidance response to the wrong word. Since this response is weak, it would, according to the hypothesis, be forgotten rapidly. But if it were, the tendency to say the wrong word would gain in potency (since the tendency to avoid it has declined) and again overall forgetting and not reminiscence would result. Once more application of the differential rate principle leads to a prediction contradicted by fact.

(2) There is the possibility that reminiscence in discrimination learning and in serial learning do not have the same explanation. The differential rate hypothesis is salvaged to the extent that it may still apply in the limited field of serial learning. Aside from the general unlikehood of there being principles of such limited scope, there is one important reason why this possibility is not probable. Both discrimination and serial learning functioned similarly in similar situations. That is, in both techniques the same increases (from 2 to 4 seconds) in the presentation intervals greatly reduced total learning time and did away with reminiscence after a 2-minute rest. Since the two kinds of learning reacted so similarly to an increase in the speed of presentation, it is likely that they are operating according to the same laws.

(3) If the first conclusion is accepted as highly probable and the second as reasonably so, there remains the third possibility that the differential rate of forgetting hypothesis is not the correct explanation of any kind of reminiscence. At present this seems the most likely alternative. But if the traditional explanation of reminiscence is not acceptable, what does account for the phenomenon?

The other possible theories have been considered one by one by Hovland (**4, 5, 6, 7**) and eliminated by him in favor of the differential rate of forgetting. The one other theory of any importance that has remained in the running explains reminiscence as being due to the dissipation with rest of a performance decrement which has accumulated during original learning and obscured the true amount learned. Such a theory is supported in particular by the fact that the absolute *sine qua non* for the production of reminiscence has been a forced rate of performing, a massing or crowding (**2**) of the material to be learned. It is quite likely that under such pressure performance would lag behind learning until a rest period gave the subject an

primary response strength would be more than compensated for by the disproportionate loss in the interfering generalized response [*cf.* speed of extinction of generalized responses (3, p. 183)]. This is a rather different form of the differential rate of forgetting hypothesis but it should be considered a theoretical possibility (if not probability) alternative to the performance interpretation adopted here.

opportunity to catch up, figuratively speaking. The major drawback of such a theory so far has been its incompleteness. No one has attempted to analyze in detail the factors that might produce such a decrement or the conditions that might dissipate it during the rest period. In an effort to meet this need and to fill the gap left by the apparent improbability of the differential rate of forgetting theory, the author has outlined a first approximation of a performance decrement theory in which the decrement is caused by conflict between responses involved in performing the task. In the present instance the decrease in presentation time increases the conflict between the confirming reaction and the choice reaction. That is, the subject at the 2-second rate is forced to confirm his last choice at the very moment when he must be making his next choice. Thus in his hurry he may easily make mistakes that he would not otherwise have made. His performance does not truly represent his learning. Rest enables performance to catch up [8]—a gain which appears as an increase in score or reminiscence. A more detailed presentation of this theory will appear in subsequent articles in this series.

The secondary significance of these data is that they duplicate Hovland's reminiscence results with the 2- and 4-second rate of presentation (**5**), but with materials that preclude the formation of remote associations. This means that in his experiment with paired associates (**7**) it is not likely that it is the lack of remote associations that is responsible for his failure to obtain reminiscence. That is, if it were lack of remote associations that was responsible for his failure, a similar lack in discrimination learning ought to have prevented the appearance of reminiscence likewise. Since this was not the case, it is probable (unless the two events have different explanations) that Hovland's failure to find reminiscence in paired associate learning is due to some other factor.

This other factor may well be the length of the presentation interval that he used. He had found with serial learning that a 4-second presentation interval destroyed the reminiscence effect (**5**) obtained with the 2-second rate of presentation. He reasoned that the crucial factor was the difference in the *response time*. So in the paired associate experiment he made the response time comparable to the 2-second rate of responding required in serial learning. However, the total presentation interval was 4 seconds. The stimulus member of each pair appeared for 4 seconds. During the first 2 seconds of this period the subject was to make his anticipation. During the last 2 seconds the response member of the pair appeared giving the subject time to check his response.

[8] This is obviously an inadequate statement. An attempt will be made later to show how rest may reduce a performance decrement. For the present it must suffice to refer for confirmation to analogous recovery with rest from work decrement (**19**) and adaptation (3, p. 107).

It is quite possible—as Hovland recognized—that this makes the recitation procedure more like the 4-second rate of serial presentation for which no reminiscence was obtained. At the 2-second rate in serial learning the subject had both to check his last anticipation and make his next one all in 2 seconds. In the paired associate experiment he had 2 seconds for each response. Consequently there was much less possibility of the two responses coming in conflict. According to the present theory as to the origin of reminiscence, if there is no conflict performance gives such a true picture of learning that rest will make little difference. So Hovland should not have expected reminiscence.

This analysis is rendered plausible by the fact that the subjects in this experiment who learned at a 4-second rate likewise had only 2 seconds to respond. The added 2 seconds permitted the subject to check his response by listening for the bell and determining whether he was right or not. So at this rate there could be no conflict between the choice reaction and the checking (or confirming) reaction such as undoubtedly existed at the 2-second rate when the reward interval (time allotted for the checking reaction) was cut from 2 to .67 seconds. In both paired associate and discrimination learning it is the cutting down of the total presentation interval which, by increasing conflict between responses the subject must make, causes reminiscence. If Hovland were to reduce his response-word presentation time considerably, as was done here, it would be expected that he would obtain reminiscence in paired associate learning.

In all serial verbal learning no matter how carefully the interval is filled with color-naming or any other activity, there are always pauses of several seconds at the end of learning and before it begins again. These periods may be used for rehearsal and no matter how explicit the instructions are not to use them for this purpose, there is always the possibility that they may be. Serial lists are particularly adapted to rehearsal because groups of words to the limit of the subject's memory span may be abstracted from the beginning and end of the list and gone over between trials. Hence the validity of the whole phenomenon of reminiscence may be questioned so long as it is found only with serial materials. Woodworth is of this opinion (**18**). Hovland's paired associate experiment (**7**) might lend support to such a contention since it was non-serial and no reminiscence was obtained. The present data, however, render further adherence to such an explanation of the phenomenon more difficult than ever. The amount of rehearsal possible when randomization of presentation is used is appreciably less than in serial learning as grouping from the beginning and end is not possible. Yet reminiscence appeared under this condition.

Summary and Conclusions

96 subjects learned 20-pair serial verbal discrimination lists on two separate days, the first for practice. 48 subjects (groups A and B) learned at a 4-second rate of presentation and 48 (groups C and D) at a 2-second rate. For 24 subjects of each 48 (groups B and D) a rest period of 2 minutes for a steadiness test was interpolated after they had reached a criterion of 15 correct choices out of a possible 20. The other 24 subjects of each 48 (groups A and C) continued learning as before to mastery without interruption. The following conclusions were reached.

(1) Discrimination learning at a controlled rate of presentation is a useful additional tool for studies of learning and memory.

(2) Learning by this method was significantly harder at a 2- than a 4-second rate of presentation.

(3) No reminiscence was obtained after a 2-minute rest if the presentation rate was 4 seconds. If the rate was 2 seconds, reliable reminiscence was obtained as measured by relearning scores.

(4) The reminiscence obtained cannot readily be explained according to the usual differential rate of forgetting hypothesis because there is no acquired wrong response in discrimination learning which by being forgotten rapidly with rest could produce reminiscence.

(5) An alternative hypothesis which will account for the results has been tentatively suggested. It maintains that reminiscence is due to a recovery with rest from an inhibition which results from conflict between recitation responses required, and which ordinarily keeps performance below the level of true learning. In discrimination learning the recitation responses in conflict are thought to be the choice reaction and the confirming reaction both of which must be made in the same short time interval with the 2-second rate of presentation.

(6) Reminiscence was obtained when remote associations were not present, a fact which suggests that Hovland's failure to find reminiscence with paired associates (**5**) ought to be reinterpreted. One such interpretation attributes his failure to the length of the presentation interval which, like the 4-second interval in serial learning, was so long that there was not sufficient conflict generated for reduction in conflict by rest to matter.

(7) Reminiscence was obtained when opportunities for rehearsal were further minimized by randomized presentation of the pairs to be learned.

(Manuscript received November 19, 1941)

References

1. Ballard, P. B. Oblivescence and reminiscence. *Brit. J. Psychol. Monogr. Suppl.*, 1913, 1, 1–82.
2. Buxton, C. E., & Newman, E. B. The forgetting of 'crowded' and 'isolated' materials. *J. exp. Psychol.*, 1940, **26**, 180–198.
3. Hilgard, E. R., & Marquis, D. G. *Conditioning and learning.* New York: Appleton-Century, 1940.
4. Hovland, C. I. Experimental studies in rote-learning theory. I. Reminiscence following learning by massed and distributed practice. *J. exp. Psychol.*, 1938, **22**, 201–224.
5. Hovland, C. I. Experimental studies in rote-learning theory. II. Reminiscence with varying speeds of syllable presentation. *J. exp. Psychol.*, 1938, **22**, 338–353.
6. Hovland, C. I. Experimental studies in rote-learning theory. III. Distribution of practice with varying speeds of syllable presentation. *J. exp. Psychol.*, 1938, **23**, 172–190.
7. Hovland, C. I. Experimental studies in rote-learning theory. IV. Comparison of reminiscence in serial and paired-associate learning. *J. exp. Psychol.*, 1939, **24**, 466–484.
8. Hull, C. L. The problem of stimulus equivalence in behavior theory. *Psychol. Rev.*, 1939, **46**, 9–30.
9. Hull, C. L. Hovland, C. I. Ross, R. T., Hall, M., Perkins, D. T., & Fitch, F. *Mathematico-deductive theory of rote-learning—a study in scientific methodology.* New Haven: Yale Univ. Press, 1940.
10. Humphreys, L. G. The factor of time in pursuit rotor learning. *J. Psychol.*, 1936, **3**, 429–436.
11. Krueger, W. C. F. The effect of overlearning on retention. *J. exp. Psychol.*, 1929, **12**, 71–78.
12. Luh, C. W. The conditions of retention. *Psychol. Monogr.*, 1922, **31**, No. 142.
13. McGeoch, G. O. The conditions of reminiscence. *Amer. J. Psychol.*, 1935, **47**, 65–89.
14. McGeoch, J. A. Learning. Chapter 13 *in* Boring, E. G., Langfeld, H. S., & Weld, H. P. Psychology. New York: Wiley, 1935.
15. Peterson, J. The backward elimination of errors in maze learning. *J. exp. Psychol.*, 1920, **3**, 257–280.
16. Van Tilborg, P. W. The retention of mental and finger maze habits. *J. exp. Psychol.*, 1936, **19**, 334–341.
17. Ward, L. B. Reminiscence and rote-learning. *Psychol. Monogr.*, 1937, **49**, pp. 64.
18. Woodworth, R. S. *Experimental psychology.* New York: Holt, 1938, pp. 63–68.
19. Yochelson, S. Effects of rest-pauses on work decrement. Unpublished Ph.D. thesis, Yale Univ., 1930.
20. Youtz, A. C. An experimental evaluation of Jost's laws. *Psychol. Monogr.*, 1941, **53**, No. 238, pp. 45.

EXPLORATORY DRIVE AND SECONDARY REINFORCEMENT IN THE ACQUISITION AND EXTINCTION OF A SIMPLE RUNNING RESPONSE *

BY FREDERICK A. MOTE, JR.

University of Connecticut

AND

FRANK W. FINGER

University of Virginia

INTRODUCTION

During the past several years a coördinated experimental program attacking the problem of simple animal learning has been under way in the psychological laboratories of Brown University. These studies have centered about changes observable in a simple running response of the rat (so-called operant or instrumental conditioning) under various controlled conditions of reinforcement and non-reinforcement. The situation requires the rat to traverse a simple short runway to obtain food, with the latent period of response ordinarily used as a measure of response strength. The specific phenomena demonstrated in detail have included simple acquisition, extinction, spontaneous recovery, external inhibition, disinhibition, retention, and discrimination (**1**, **3**, **4**, **5**, **6**, **10**). The effects of emotional factors, varying amounts and patterns of reinforcement, and temporal relationships, have in addition been investigated (**2**, **7**, **9**). Underlying the group of studies has been the hope that the results, obtained under closely comparable conditions, may eventually be integrated in terms of a quantitative hypothesis of the learning process (*cf.* esp. **6**).

Implicit in most of these studies have been 2 assumptions: (1) that the progressive changes in response occurring during acquisition and extinction are the manifestation of the changing strength of the learned correlation between conditioned stimulus (the pattern of stimulation arising from the starting box, and particularly the opening of the starting-box door) and unconditioned stimulus (food); and (2) that the only significant reinforcement in the situation is the administration of the food pellet at the termination of a trial. In

* This experiment was conducted in the Psychological Laboratory of Brown University.

other words, the problem has been conceived as involving an uncomplicated instance of learning, with reinforcement added to or withdrawn from the situation by the manipulation of food pellets. The significance of these assumptions varies with the precise nature of the problem, but they are of utmost importance in any theoretical formulation of the mechanism underlying the behavior changes. It was the purpose of the present experiment to obtain evidence bearing on the validity of these suppositions.

In examining the first assumption, the questions were raised: Is there any progressive change in the running response during a series of trials none of which are reinforced by feeding? Would the orderly changes in latent time which have been shown to occur during acquisition have taken place had no reinforcement been given? To what extent is this demonstrated trend the reflection of some such complicating factor as an 'exploratory' drive? The test of these points involved the running of the rats in a situation corresponding to the usual acquisition situation, with the exception that the reinforcing agent (actual feeding, or access to a familiar food box) was omitted. It will be shown that in spite of lack of apparent objective reinforcement, the animals continue to respond for several trials. The relatively slow and uniform response time, however, suggests that a simple correction can be applied in the usual situation involving reinforcement, so that the course of the learning process itself may be followed with sufficient assurance of reliability.

The testing of the second assumption may be paraphrased in the query: Is the only reinforcing agent operating in the situation the actual obtaining of food at the end of the runway? In other words, can entrance into a goal box already associated with feeding produce some positive effect upon subsequent response, even though the actual feeding is lacking? The answer to these questions was furnished by a comparison of results obtained when no food was given but a familiar goal box was used, with those found when neither food nor the familiar box was provided. The analysis of the data will indicate that the familiar goal box is in itself a source of secondary reinforcement.

Apparatus

The apparatus used in this experiment was a slightly modified form of that devised by Graham and Gagné (6). It consisted essentially of a simple elevated runway, 3′ long and $1\frac{1}{4}''$ wide, at either end of which was a small platform. On each platform could be placed a box ($5\frac{1}{2}'' \times 6'' \times 7''$), open in front and covered on top with wire mesh. Four such boxes were provided (designated, respectively, A, B, C, and D), 2 of which were used alternately as starting box and goal box during a given experimental session. Between the end of the runway and each platform was a wall-and-sliding-door arrangement so constructed that the rat could be released or detained in the box according to the experimental schedule. The whole apparatus was contained within a wooden frame, covered with cheesecloth. Since the only illumination came from a 25-watt

lamp within the framework (suspended 12 inches above the runway and 8 inches in front of the goal box), the cloth provided an effective one-way screen. There was also in the experimental room an 8-foot long, 1-inch wide practice runway, on which preliminary training was given.

Procedure

The experimental routine was designed to duplicate that typically used in this learning situation, with 2 important deviations: in one group of animals the food reward was omitted, and in a second group food was omitted and in addition an unfamiliar goal box was used.

The preliminary training of Group I (Secondary Reinforcement) was practically identical with that given in previous experiments. On the first day the 24-hour-hungry rat was placed on the practice runway and allowed to run about for 10 minutes. After about 1 minute's wait on a small platform he was permitted to enter box A or B, and placed in the starting-box position of the apparatus for 5 minutes. The box was then transferred to the food-box position and a 20-minute feeding period given. Three similar training sessions followed at 24-hour intervals; only a 15-minute feeding period was given on the last of these days. On the experimental day the rat was placed on the practice runway for 2 minutes and then allowed to enter box A which was immediately placed in the starting position. After 3 minutes the door was opened and a stop watch started. When the rat crossed a mark 4 inches down the runway this watch was stopped and a second watch started. The first watch thus indicated the length of the latent period, which in the previous studies has been assumed to vary inversely as the strength of the learned response. When the rat entered box B, in the goal position, the doors were closed and the second watch was stopped, yielding the measure called 'running time.'[1] At the same instant a third watch was started, allowing the timing of the inter-trial interval. After the rat had been in the (empty) box B for 30 seconds, the box was removed from the goal position and placed in the starting position, while box A replaced B in the goal position. After an inter-trial interval of 75″ (with the rat approximately 30″ in the goal position and 45″ in the starting position) the starting door was again opened and the second trial begun. This procedure was followed until that trial in which the animal did not leave the starting box within 5 minutes, the criterion selected to indicate 'complete extinction.'[2] At the end of this 'extinction' interval the door was reclosed. During a period of either 4½ or 17 minutes the rat was left in the closed starting box. After this 'spontaneous recovery' interval the door was again opened and the animal once more run to the criterion of extinction.[3]

The treatment of the Group II animals (No Reinforcement) differed from this only in that the feeding during the preliminary sessions was in no way connected with the experimental apparatus. After each of the preliminary periods on the practice runway the rat of this group was placed in box C or D in the starting position for 5 minutes. He was then kept in a neutral box for 1 hour, after which he was fed for 20 (or, on the fourth day, 15) minutes in still another box. On the experimental day he was run in exactly the same manner as were the animals of Group I, boxes A and B being used interchangeably. Thus, the rat had had acquaintance, prior to the experimental session, with neither the boxes used (A and B) nor the goal position of the

---

[1] The use of this supplementary measure was suggested by Verplanck's study (10), in which the running time proved in certain instances a more differentiating measure than latent period. The results of the present experiment confirm the advisability of obtaining this characteristic of the rats' response. It appears that differential experimental treatment may sometimes be reflected in the response in terms of running time differences, while the latent period reveals no significant difference.

Whenever a rat retraced from past the 4-inch mark, to the starting box, the running-time watch was stopped and the latent-period watch restarted, so that the final measures given represent the total time spent during the complete trial in the respective parts of the apparatus.

[2] In the previous studies the 3-minute criterion was used; the data corresponding to this criterion are also available in the present records.

[3] Two different spontaneous recovery intervals were used in order to check the applicability of Graham and Gagné's theory (6) to this aspect of our situation. Some of the previous learning studies employing this apparatus have found, confirming the prediction based on the theory, that the longer spontaneous recovery interval led to greater spontaneous recovery. Our Group I data do not confirm this supposition, although the figures for Group II tend in the expected direction.

apparatus: his feeding had never been associated (at least, within the temporal limit of 1 hour) with the situation employed during the experimental trials.

The nature of the procedure may be summarized as follows: The animals of Group I were treated in essentially the same manner as have been the rats in earlier experiments employing this apparatus, with the exception that reinforcement by food was never given at the end of the successive trials. Whatever differences from previous results may be observed in their performance may therefore be attributed to this lack of reinforcement. The treatment of the Group II animals differed from that of Group I primarily in that the possibility of secondary reinforcement at the end of the various trials (in the form of entrance into a situation previously associated with feeding) was obviated. Any differences in response between the two groups are thus dependent upon the respective presence or absence of secondary reward.

The Animals

Each group was composed of 15 male albino rats from the laboratory colony. All 30 rats belonged to a single litter-group (siblings and double first cousins), and averaged 74 days of age (range 67 to 93 days) on the experimental day. The animals were selected randomly and tested in such a sequence that the 2 groups were perfectly matched with respect to age. Thirty-eight rats were given the preliminary training, but 8 were discarded for failure to respond appropriately during the experimental session.

Results

Group I.—The data for Group I (Secondary Reinforcement) are presented in Tables I and II and Figs. 1 and 2. Because individual results are extremely variable, averages are used to represent the group trends. It may be seen in Table I that the individual rats

TABLE I

Average Data, and Comparison of Groups I and II

Measure	Group I (Sec. Reinf.)	Group II (No Reinf.)	Difference	*t*	*P*
Number of trials to 5-minute criterion	27.7	13.3	14.4	1.8	9%
Range of trials to 5-minute criterion	6−99=93	2−42=40	53		
Number of trials to 3-minute criterion	16.4	8.3	8.1	1.2	25%
Range of trials to 3-minute criterion	1−99=98	1−25=24	74		
Log latent period of trials 1–10	1.61	1.70	0.09	1.5	15%
Log running time of trials 1–10	1.09	1.17	0.08	1.3	20%
Log running time of trials 1–5	1.02	1.22	0.20	3.3	1%
Number of spontaneous recovery trials	11.3	2.7	8.6	2.1	<5%
Log latent period of first spontaneous recovery trial	1.99	2.07	0.08	0.6	55%
Log running time of first spontaneous recovery trial	0.95	1.41	0.46	2.7	1%
Log latent period of all spontaneous recovery trials	1.87	1.97	0.10	1.0	30%
Log running time of all spontaneous recovery trials	1.09	1.41	0.32	3.2	<1%

TABLE II

AVERAGE LATENT PERIOD AND RUNNING TIME FOR TRIALS 1–13

Trial	Group I					Group II				
	Lat. Per. (Seconds)	Log L.P.	Run. Time (Seconds)	Log R.T.	N	Lat. Per. (Seconds)	Log L.P.	Run. Time (Seconds)	Log R.T.	N
1	74.1	1.87	14.1	1.15	15	83.2	1.92	16.2	1.21	15
2	28.2	1.45	7.8	0.89	15	32.4	1.51	12.6	1.10	15
3	31.6	1.50	10.0	1.00	15	35.5	1.55	15.1	1.18	14
4	38.0	1.58	9.8	0.99	15	45.7	1.66	18.2	1.26	13
5	38.9	1.59	12.0	1.08	15	38.0	1.58	22.9	1.36	12
6	38.9	1.59	9.1	0.96	15	47.9	1.68	18.6	1.27	11
7	42.7	1.63	11.2	1.05	13	44.7	1.65	14.5	1.16	10
8	51.3	1.71	17.8	1.25	12	51.3	1.71	7.6	0.88	9
9	38.0	1.58	20.4	1.31	11	83.2	1.92	14.8	1.17	9
10	42.7	1.63	17.0	1.23	10	70.8	1.85	11.5	1.06	8
11	74.1	1.87	17.4	1.24	9	74.1	1.87	11.5	1.06	7
12	40.7	1.61	20.0	1.30	8	60.3	1.78	7.4	0.87	6
13	51.3	1.71	17.4	1.24	8	58.9	1.77	19.5	1.29	6

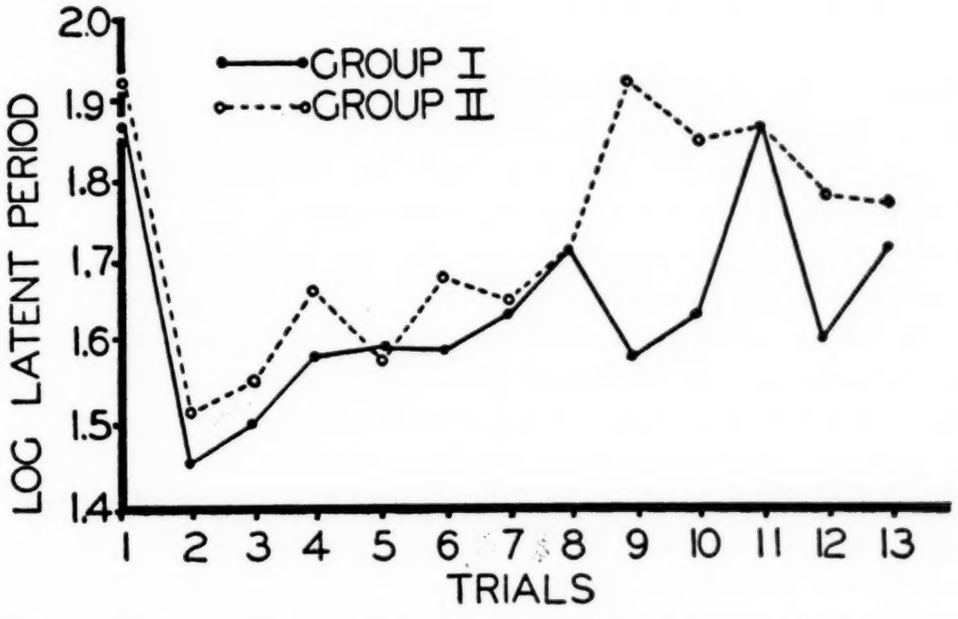

FIG. 1. Mean log latent period for Groups I and II, trials 1–13 (see Table II).

required between 6 and 99 trials to reach the 5-minute criterion of extinction, with an average of 27.7 trials. The mean log latent period of the first 10 trials was 1.61, and the mean log running time for the same trials 1.09. The mean log latent period on the first trial of spontaneous recovery was 1.99,[4] with the mean log running

[4] If an animal failed to respond within the 5-minute criterion on the first spontaneous recovery trial, a log latent period of 2.50 (the approximate value of 300 seconds) was arbitrarily assigned to him. This figure is of course too small, since it is certain that the latent period of any eventual response would have been considerably greater than this figure. This fact, however, only increases the significance of the difference later pointed out between the 2 groups.

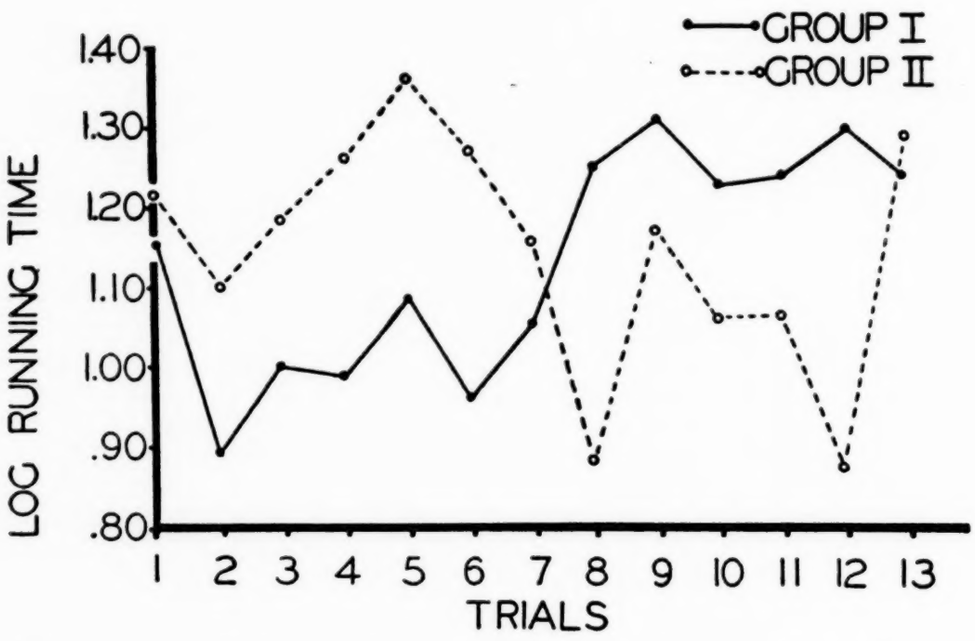

FIG. 2. Mean log running time for Groups I and II, trials 1–13 (see Table II).

time 0.95. The average number of spontaneous recovery trials required to reach the 5-minute non-response criterion was 11.3. The mean log latent period and log running time for all spontaneous recovery trials were 1.87 and 1.09, respectively.

The average curve representing the log latent period of the animals' responses is given in Fig. 1 (solid line), based on the values from Table II. The number of animals contributing to the mean points varies, because of the different extinctive resistance of the several animals. All 15 rats' times are included in the first 6 points, while only 8 animals are represented on the 13th trial. The curve begins at 1.87 log units, drops significantly to 1.45 on the second trial ($t = 3.3$), and then rises slightly for 2 trials. Subsequent points deviate relatively little from 1.62 log units. The log running time curve (Fig. 2, solid line) is somewhat more variable. The first half of the curve remains quite uniformly near 1.00 log unit, while the last 6 points cluster around 1.25 log units.

Group II.—The results of Group II (No Reinforcement) are presented in Tables I and II and Figs. 1 and 2. In Table I it is shown that the range of trials required to reach the 5-minute criterion of extinction was from 2 to 42, with an average of 13.3. The mean log latent period of the first 10 trials was 1.70, the mean log running time 1.17. The first trial of spontaneous recovery had a mean log latent period of 2.07, while the running time figure was 1.41 log units. An average of 2.7 spontaneous recovery trials was required to bring the animals to the second 5-minute criterion of non-response. The mean

of all spontaneous recovery latencies was 1.97 log units, while the corresponding log running time value was 1.41.

A comparison of the data for the 2 groups is also given in Table I. Group II required 14.4 fewer trials to reach the 5-minute criterion of non-response (significant at the 9 percent level) and 8.1 fewer to reach the 3-minute criterion. The average Group II latent period for trials 1–10 was 0.09 log units longer than that of Group I, running time 0.08 longer for trials 1–10 and 0.20 longer for trials 1–5 (significant at the 1 percent level). Both the mean latent period and the running time of the first spontaneous recovery trial were longer in Group II, the difference of the latter being significant at approximately the 1 percent level. It required 8.6 fewer spontaneous recovery trials for Group II to reach the criterion of extinction ($P < 5$ percent) and the mean running time of all spontaneous recovery trials was 0.32 log units greater (significant beyond the 1 percent level).

The curves of the average Group II data (Figs. 1 and 2, broken lines; *cf.* Table II) are somewhat different from those of Group I. The various points of the log latent period curve, based on 15 cases for the first 2 trials and as few as 6 cases by the 13th trial, are with one exception higher than the corresponding values on the Group I curve (see Fig. 1 and Table II). The distinguishing feature of the curve is its centering around 2 levels—about 1.66 log units for the first 9 trials, and 1.82 for the remaining trials. The significance of these tendencies is questionable, since the number of cases is so variable at the successive points. The same criticism applies to the log running time curve (Fig. 2). This curve as a whole is more variable than that of Group I; it is higher than the Group I curve in its first portion, and then lower for the remaining trials.

The comparison of the data for the 2 groups may be summarized by the statement that the Group I rats (Secondary Reinforcement) responded with the shorter latent period and running time, continued to respond for a greater number of trials, and showed more prompt and longer lasting spontaneous recovery.

Discussion

The data which this experiment yields would seem to provide sufficiently definite answers to the questions originally propounded, clarifying, rather than invalidating, the results of previous studies involving this learning situation. In the first place, certain progressive changes in the running response do occur when no apparent reinforcement is given—some exploratory drive does impel the animal to respond even though controlled reward is withheld. That measure

which has been assumed to be a direct indicator of strength of learned response now proves to be complicated by the operation of another factor. In the second place, it rather clearly appears that the administration of food is not the only reinforcing agent that can affect the response: the entrance into the familiar feeding box at the end of a run exerts in itself a positive influence on later trials.

If it is thus shown that the response trend previously assumed to represent a change in the learned connection between conditioned stimulus and food, is in part the manifestation of an exploratory drive, the value of this situation for the study of learning is called into question. Since a significant decrease in latent period occurs between the first and second trials when no reinforcement is given at the end of trial 1, should the drop after reinforcement be interpreted as an indication of learning? A closer examination dispels the doubts raised by this question. The non-reinforced rats did respond faster on the second trial than on the first, but this drop was in the range of only 0.44 log units, as compared to a mean drop as great as 0.93 log units with reinforced rats (*cf.* **2**, **6**). The reinforcement of the first trial, then, produces a greater decrease than can be accounted for in terms of exploratory drive; that is, there is an additional effect attributable to learning.[5] When the comparison is extended beyond the second trial a still more apparent effect of reinforcement (*i.e.*, learning) may be seen. The average latent period of the non-reinforced animals *increased*, while the corresponding times of reinforced animals in the earlier experiments continued to decrease in an orderly fashion. In order that a comparison can be made, in Fig. 3 are presented the mean log latent period curves for Groups I and II of this study, along with the curve of acquisition obtained by Graham and Gagné (**6**, p. 260) for rats receiving reinforcement on every trial. The lowest point in the latent period curve of the 'exploring' animals is 1.44, while rats 'learn' to respond in a latent period of about 0.50 log units (12 to 16 acquisition trials; *cf.* **2**, **6**, and **9**). There is no question, then, but that in the studies of learning utilizing this situation, a regular change in behavior occurs that must be attributed to something more than a drive to explore.

It may thus be agreed that the definite trend in behavior which takes place with appropriate reinforcement is due in part to learning. The objection may be raised, however, that the exact measurement of this function is complicated by the operation of the additional factor of exploratory drive. To what extent is a particular latent period the reflection of learning strength, and to what extent the

[5] The average latent period drop from trial 1 to trial 2 of the Group I rats was 0.49 log units less than the latent period drop of 15 animals randomly selected from Finger's experiment (**2**). This difference is significant far beyond the 1 percent level ($t = 37.6$).

manifestation of drive to explore? This point is perhaps pertinent; it can be resolved, however, without too much difficulty. It is seen in Fig. 1 that, after the first trial, the average log latent period of Group I varies little from trial to trial, ranging rather closely about 1.60. The exploratory drive has a fairly constant value, at least between the 2nd and the 10th trials. It is a simple matter to apply a constant correction to the typically obtained learning curve, regarding the changes in latent period as variations from the exploratory base of approximately 1.60 log units. Any drop in response time below this base would then be considered a relatively accurate

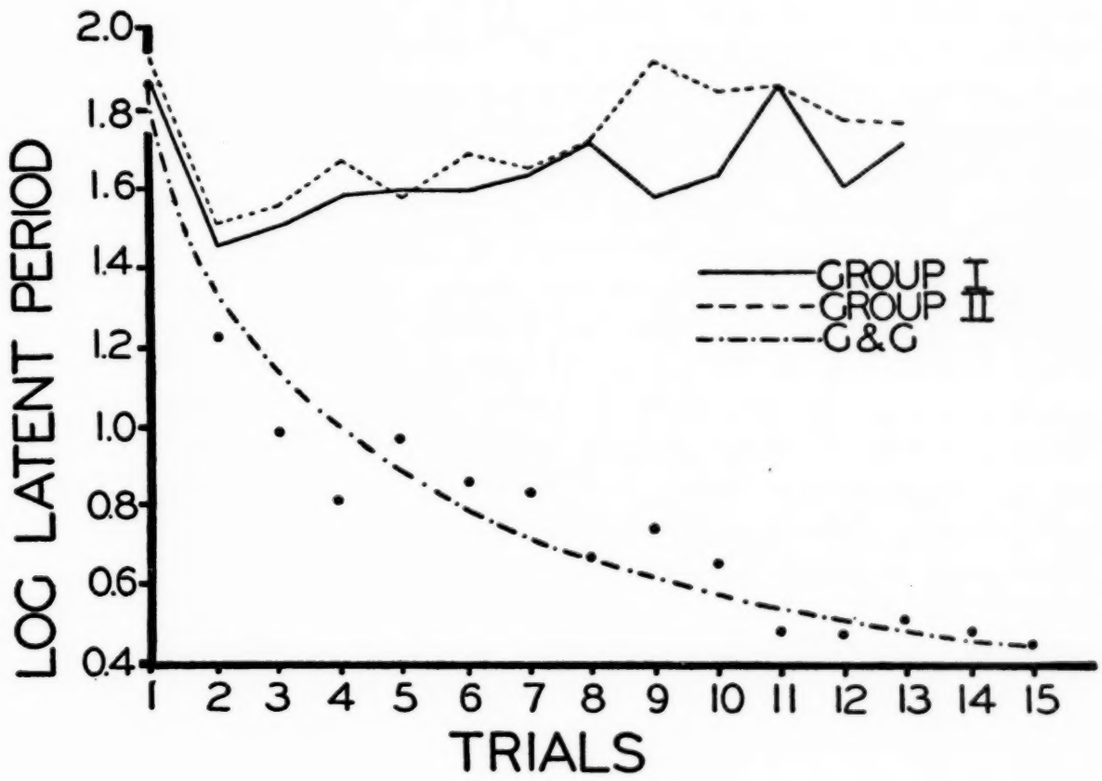

FIG. 3. Comparison of mean log latent period curves of this study for Groups I and II (no food reinforcement) and curve of acquisition (food reinforcement every trial) obtained by Graham and Gagné (6).

indication of the progress of learning. As far as the first point in the learning curve is concerned, the present data tend to corroborate, rather than refute the original theoretical formulation of Graham and Gagné. It will be recalled that in attempting to fit their empirical data to their theoretically derived curve representing the course of acquisition, a hypothetical 'true' beginning of learning was extrapolated backward from the latter part of the curve (6, p. 274). This derived first point falls, in their study, at about 1.40 log units. The corresponding value (obtained by visual fit) for Finger's data (2) is

about 1.50.[6] It is perhaps more than a coincidence that this latter figure is in the same range as that representing the quite uniform exploratory level of the animals in the present experiment. The point from which, on the basis of theoretical computation, Graham and Gagné trace the 'true' course of learning is not far removed from the point which we would establish as the base from which the learning trend should be followed. The point fulfilling the terms of their quantitative theory may thus be confirmed by our obtained data.

The second conclusion which may be drawn from the results is that some reinforcing value is attached to the situation in addition to the usual reward factor (food). To be sure, Hilgard and Marquis' statement that reinforcement can be "identified by the fact that when it is present on successive training trials the conditioned response increases in strength" (**8**, p. 69) is not unequivocally fulfilled. The strength of the secondary reinforcement is insufficient to produce a learning increment detectable by a latent-period decrease,[7] but the fact that every possible comparison of Groups I and II points to a stronger response-tendency in the Group I animals leads us to believe that there was a positive factor at work in the Group I situation which was absent in the Group II situation. The only difference in conditions can be described in terms of familiarity with the goal boxes in Group I: these animals were 'rewarded' at the end of each run by surroundings which had been associated with feeding in the past (*cf.* **11**). The latent period measure reveals the operation of this factor only in relative terms (Group I responds more promptly than does Group II), but to the extent that other indices (*e.g.*, number of trials for extinction, the latency of the first spontaneous recovery trial, the number of spontaneous recovery trials) reflect 'strength' of the response the existence of secondary or sub-goal reinforcement is indicated.

The operation of secondary reinforcement in this learning situation is a complicating factor which must be considered whenever the nature or amount of reinforcement is subject to investigation. Thus, when the reinforcing value of food is considered in relation to its spread through time (**4**), or when the effect of varying amounts of reinforcement given in different patterns is studied (**2**, **9**), the secondary reinforcement value of the goal box becomes a relevant factor. This, coupled with the continued operation of 'exploration' for at

[6] This figure is more appropriate for comparison with the present study, because of the closer approximation of age means of the rats.

[7] It might be expected that the strength of the secondary reinforcement would decrease as the goal box is presented successively without the primary reinforcement of food. In other words, the connection between goal box and food should itself become gradually extinguished (*cf.* **11**, p. 48).

least 16 trials (3-minute criterion), would be expected to complicate the course of normal extinction. These facts would seem to clarify certain anomalous results previously reported,[8] while requiring further consideration of data before the final quantitative formulation can be attained.

Summary

Two groups of 15 rats each were tested in a situation involving a simple running response. Latent period of reaction and running time of each trial were measured, as well as number of trials required to bring the animals to an extinction criterion of 5 minutes of non-response, and the extent of spontaneous recovery. The following conditions were used:

Group I—Reinforcement by food at the end of each trial was omitted, but the goal box used was one previously associated with feeding;

Group II—Neither food nor the familiar goal box was given during the experimental trials.

1. In spite of lack of reinforcement by food, the Group I animals responded for an average of 27.7 trials, with mean log latent period on the first 10 trials 1.61 and log running time 1.09.

2. After the first trial, the Group I curve of mean log latent period maintained a fairly uniform level for 10 trials.

3. Every comparison of Groups I and II (average latent period and running time, number of extinction trials, latent period and running time of the first spontaneous recovery trial, number of spontaneous recovery trials, mean latent period and running time of all spontaneous recovery trials) pointed toward the existence of a stronger response tendency in the Group I rats.

The results indicate that the rats are impelled by some exploratory drive to make the running response even if reinforcement is never experienced in the situation. This factor may complicate the quantitative specification of learning by requiring consideration of an additional parameter. The fact that Group I responds more strongly

[8] That non-reinforced rats continue to run for an average of more than 16 trials, purely on the basis of exploratory drive, may explain why rats with 6, 8, or 12 trials during the acquisition series continue to respond as long (or longer) during extinction as do those with a longer acquisition series (**2, 9**). As was suggested in an earlier paper, the exploratory drive continues to operate during the early part of extinction, exerting a stronger effect the shorter the acquisition experience and thus perhaps lengthening the extinction series (**2**). The functioning of secondary reinforcement also fits the earlier suggestion that attempted to explain why a 16-trial, 8-reinforcement acquisition series should lead to a slightly greater resistance to extinction than does a series consisting of 8 successive reinforced trials. It appeared that the 8 acquisition trials during which no reinforcement by food was given had a positive effect, revealed during extinction. This positive effect, it now seems, may be attributable to the secondary reinforcement derived from the familiar food box.

than Group II is interpreted as an indication that the presence of a familiar goal box may provide a significant amount of secondary or sub-goal reinforcement.

These conclusions are considered with respect to previous learning experiments utilizing this situation, with reference to their significance for a quantitative description of the learning process.

(Manuscript received November 22, 1941)

References

1. Berry, R. N., Verplanck, W. S., & Graham, C. H. The reversal of discrimination in a simple locomotor habit. (To appear.)
2. Finger, F. W. The effect of varying conditions of reinforcement upon a simple running response. *J. exp. Psychol.*, 1942, **30**, 53–68.
3. Gagné, R. M. External inhibition and disinhibition in a conditioned operant response. *J. exp. Psychol.*, 1941, **29**, 104–116.
4. Gagné, R. M. The effect of spacing of trials on the acquisition and extinction of a conditioned operant response. *J. exp. Psychol.*, 1941, **29**, 201–216.
5. Gagné, R. M. The retention of a conditioned operant response. *J. exp. Psychol.*, 1941, **29**, 296–305.
6. Graham, C. H., & Gagné, R. M. The acquisition, extinction, and spontaneous recovery of a conditioned operant response. *J. exp. Psychol.*, 1940, **26**, 251–280.
7. Graham, C. H., & Gagné, R. M. The effect of an 'emotional state' on the initial stages of acquisition in a conditioned operant response. *Proc. nat. Acad. Sci.*, 1940, **26**, 297–300.
8. Hilgard, E. R., & Marquis, D. G. *Conditioning and learning.* New York: Appleton-Century, 1940. Pp. xi + 429.
9. Mote, F. A., Jr. The effect of varying amounts of reinforcement upon the acquisition and extinction of a simple running response. (In preparation.)
10. Verplanck, W. S. The development of discrimination in a simple locomotor habit. (To appear.)
11. Williams, K. A. The reward value of a conditioned stimulus. *Univ. Calif. Publ. Psychol.*, 1929, **4**, 31–55.

THE SPECIFICITY OF THE EFFECT OF SHOCK ON THE ACQUISITION AND RETENTION OF MOTOR AND VERBAL HABITS *

BY JACK BERNARD †

From the Psychology Department of Clark University

The problem of the generality versus the specificity of the effect of electric shock administered for error in learning situations has been discussed at length in an earlier paper (1). At that time it was found that shock administered in connection with some but not all errors in a stylus maze situation resulted in significantly faster elimination of shocked than non-shocked errors.

This finding raised the following questions:

1. Was the obtained specificity a function of the particular maze pattern used in that experiment, or is it characteristic of stylus mazes in general?

2. Can it be found in learning situations other than the stylus maze?

3. Will it be evidenced in relearning trials some days after the original learning, or is it lost as learning is completed?

The importance of the concept of shock specificity depends directly upon the answers to these questions. If such specificity is observable only under certain highly-restricted conditions, it is of little more than academic interest. If it is more widely characteristic of learning under conditions of punishment for error, it may have important educational implications. The present study is intended to broaden the scope of our original investigation in an effort to answer the above questions. In order to answer the first question, we have made use of a unidirectional stylus maze in which the pattern is basically different from that of the Fox maze of the previous study. To throw light upon the second question, we have added a verbal maze adapted for mechanical presentation from that of Peterson (6). The third question has led to extending the study beyond learning to include recall and relearning trials one week later.

* A more detailed account of this study is on file at the Clark University library in the form of an M.A. thesis.

† I wish to thank Dr. Robert H. Brown for his helpful assistance throughout the course of this study.

Subjects, Apparatus and Procedure

Subjects.—The subjects employed in this study were 62 undergraduate male students at Clark University, selected at random from the undergraduate student body. They were divided into eight groups, also on the basis of chance. The task for each subject was to learn both the verbal and the motor mazes to completion, and then to relearn each after a seven-day interval. Before considering the procedure in detail, the two mazes will be described.

Motor Maze.—The design of the motor maze is illustrated in Fig. 1. The maze was constructed of 5-ply wood after the method described by Miles (5). The parts were tacked firmly to 1¼ inch birch covered with a sheet of galvanized iron. This gave a smooth surface upon which the stylus could travel.

The experimental maze consisted of 14 units, in which the pattern was LRRRLLRLLRRLLR. Metal contacts were placed at the end of each cul-de-sac and wired through the bottom of the maze to two separate electrical leads. These leads were connected with a two-way switch in such a manner that current could be sent to either the odd- or even-numbered blinds at will. The shocking current came from a Porter inductorium connected for tetanic shock and activated by

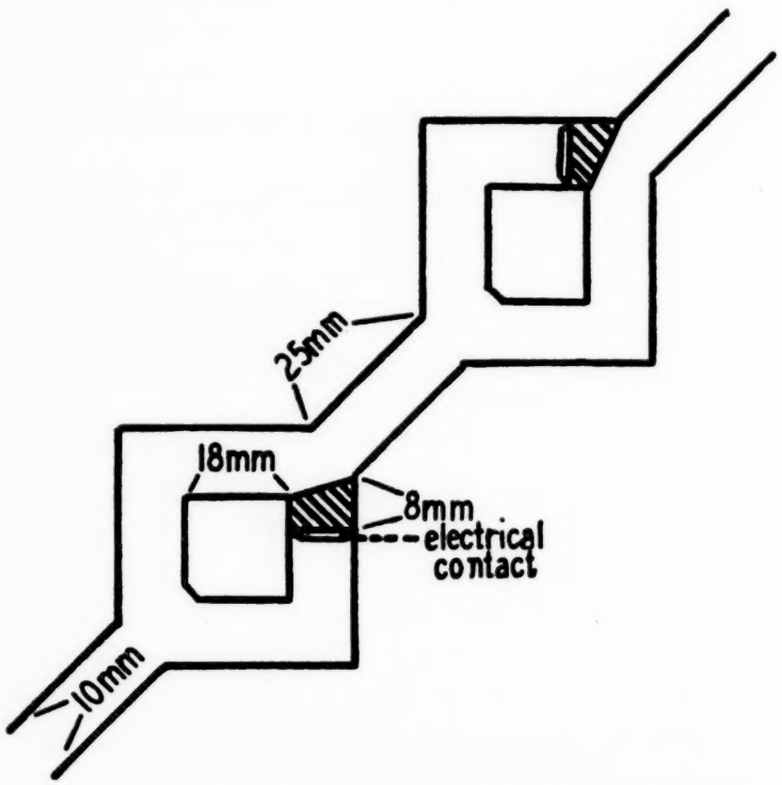

Fig. 1. Design of motor maze used in this experiment.

two 1.5-volt dry cells, and was administered directly to the active hand by means of the New York University stylus [1] (Crafts & Gilbert, 2) when the stylus touched the contact at the end of a shock blind. The units of the maze were numbered 1–14 consecutively.

The demonstration maze consisted of two parts, which were mounted side by side on a common base. One was a single unit identical with a unit of the experimental maze, with the blind in position. The other was a series of five units, exactly like the pattern of the experimental maze, but without blinds. The contact in the single unit was wired into the same circuit as the contacts of the experimental maze.

Verbal Maze.—The verbal maze consisted of a series of pairs of 3-letter nonsense syllables, in which one of each pair had been arbitrarily designated as 'right' and the other as 'wrong.' These syllables were selected from Hull's list (3) and with these restrictions:

[1] Loaned for this experiment through the courtesy of Dr. Ralph W. Gilbert.

(*a*) The mean association value was kept as low as practicable.

(*b*) The mean deviation of association values was low.

(*c*) The two syllables of each pair had equal association value.

(*d*) Odd and even pairs of syllables had a similar mean association value and mean deviation of association value.

(*e*) The syllable list as a whole was as far as possible in accord with Luh's rules for the construction of nonsense syllable lists (4).

Table I gives the list of syllables used, together with their individual association values, means, and mean deviations.

TABLE I

NONSENSE SYLLABLES OF THE VERBAL MAZE, WITH ASSOCIATION VALUES AND RELEVANT DATA

Pair	'Right'	Association Value	'Wrong'	Association Value
1	VEF	0	HAJ	0
2	NUY	13	RIX	13
3	WAB	18	VOM	18
4	YUT	5	ZIK	5
5	BOP	12	DAX	12
6	JEB	8	WUP	8
7	TOV	7	YIK	7
8	NAD	12	WEF	12
9	JOM	17	KIR	17
10	MEJ	13	VUR	13
11	FAP	5	WIB	5
12	ZAH	8	YUK	8
13	KIF	10	BEP	10
14	VAW	3	JID	3
15	MEP	8	ZOF	8
16	DUY	15	FEK	15
17	POF	15	ZIR	15
18	KUN	13	WOY	13
19	PIJ	3	MEV	3
20	ZUK	5	FOV	5

	Mean Association Value	Mean Deviation of Association Value
In entire list	9·5	4·3
Odd-numbered syllables	9·5	4·9
Even-numbered syllables	9·5	3·7

Four experimental lists were prepared from this key list, in each of which the position of the 'correct' syllable of each pair (on the left or the right) was determined by chance and then corrected to have as many correct choices on the left as on the right, and to prevent more than three consecutive choices from being on the same side. The 'left-right' pattern of the correct choices was different on each list. The four lists were typed one after the other in black capitals on glazed kymograph paper mounted as an endless belt on two horizontal kymograph drums, and presented to the subject one pair at a time by the experimenter's pressing of a lever which turned the kymograph drum one notch at a time. The pairs of syllables were so spaced that a movement of one notch moved one pair from the presentation window and brought the next into view. The end of each list was separated from the beginning of the next by a line of red asterisks. The use of four lists in this manner prevented the subjects from learning the correct choices on the basis of position.

Electric shock was administered by the experimenter's pressing of a telegraph key. The electrodes were two copper pennies, one held by the subject between the first and second joints of his bent left forefinger, and the other between the tips of his left thumb and middle finger. The current came from a Porter inductorium set for tetanic shock and operated on a single 1.5-volt dry cell.

All experimentation took place in a room in which lights were arranged to throw constant illumination on the maze or memory drum.

Procedure.—Each subject learned one of the two tasks to the criterion of two successive trials without error, three days later learned the other to the same criterion, and seven days after the original learning relearned each task to the same criterion of two successive errorless trials.

The eight subdivisions of the experimental group were intended to cover the eight possible combinations of the variables in the situations. These are as follows:

I. Motor maze presented first, then verbal maze.
  A. Shock on even-numbered blinds of motor maze.
    1. Shock on even-numbered syllable pairs of verbal maze.
    2. Shock on odd-numbered syllable pairs of verbal maze.
  B. Shock on odd-numbered blinds of motor maze.
    1. Shock on even-numbered syllable pairs of verbal maze.
    2. Shock on odd-numbered syllable pairs of verbal maze.

II. Verbal maze presented first, then motor maze.
  A. Shock on even-numbered syllable pairs of verbal maze.
    1. Shock on even-numbered blinds of motor maze.
    2. Shock on odd-numbered blinds of motor maze.
  B. Shock on odd-numbered syllable pairs of verbal maze.
    1. Shock on even-numbered blinds of motor maze.
    2. Shock on odd-numbered blinds of motor maze.

Typed instructions were handed each subject as he entered, and these instructions were augmented by such oral explanation as appeared necessary. The instructions for the motor maze were given with the demonstration maze in view. They described the nature of the task, and informed the subject that he would receive shock on contacting some, but not all, blinds.

The strength of shock appropriate for each subject was determined by the following method. With current flowing in the demonstration blind, the subject touched it a number of times successively while the experimenter gradually increased the strength of the current from an imperceptible amount. The subject was asked to report when it was 'too much,' at which point it was slightly decreased and the trials begun. From a more objective point of view, the strength of the shock was such as to evoke an overt reaction from the subject, but not of such a strength as to cause him to withdraw the stylus from the maze. If the subject appeared to become adapted to the shock during learning so that he no longer reacted overtly, the current was increased without mentioning that fact.

The subjects were permitted to practice on the demonstration maze as long as they wished before opaque goggles were placed over their eyes and the trials proper begun. Detailed records of the exact errors made and the time required for each trial were taken on forms mimeographed for that purpose.

Typed instructions for the verbal maze were handed the subject as he was seated before the apparatus with the asterisks between Lists IV and I in view. They described the nature of the task, and informed the subject that he would hear 'Wrong' for some errors and feel a shock for others. The amount of shock to be given each subject was determined by a method similar to that used in connection with the motor maze, and the exact errors recorded on a special form. Shock was not administered during any of the relearning sessions.

Results

The accumulated data were gathered into three categories for separate analysis. These were Learning, Recall (consisting of the first Relearning trial) and Relearning (consisting of the Recall trial plus the rest of Relearning). The Learning data were analyzed by six methods, the Recall by two, and the Relearning by four. A brief description of these methods follows.

Learning.—(*a*) The mean number of shocked errors per subject is compared with the mean number of non-shocked errors, and the critical ratio of the difference determined.

(*b*) The proportion of individuals making more shocked than non-shocked errors is compared with the proportion making more non-shocked than shocked errors, and the critical ratio of the difference determined. This prevents overweighting by extreme scores.

(*c*) The mean number of shocked and non-shocked errors respectively per trial are determined for each subject, and the shock and non-shock means summated separately. The shock and non-shock means of the individual means are then compared and the critical ratio of the difference determined. This results in the weighting of each individual's score by the density of his errors.

(*d*) *A total shock ratio* is computed by dividing the number of shocked errors for all subjects in the first half of the learning trials (excluding the two final errorless trials) by the number of shocked errors for all subjectsin the second half of the trials. The *total non-shock ratio* is computed similarly, and we are then able to compare two figures which represent the average *rate of elimination* of shocked and non-shocked errors respectively, sigmas being computed from *individual shock ratios*.

(*e*) The number of individuals having a higher shock than non-shock ratio is compared with that having a higher non-shock than shock ratio, and the critical ratio of the difference determined. This again prevents overweighting by extreme scores.

(*f*) Finally we are concerned with the efficiency of electric shock, as compared with the simple blocking in the motor maze or the word 'Wrong' in the verbal maze, in preventing repetition of errors on the following trial. Each record is examined to determine how many of the shocked errors are 'repeats,' that is, are preceded by similar errors in the previous trial. The sum of all such errors is divided by the number of subjects. The quotient is the mean number of repeated shocked errors per subject. The mean number of repeated non-shocked errors is computed similarly. The two figures are then compared and the critical ratio of their difference determined.

Recall.—Shocks were not administered during this part of the experiment. 'Shocked errors' therefore refer to errors which were shocked during learning.

(*a*) This is a repetition of the method explained in (*a*) under 'Learning,' now applied to the single Recall trial.

(*b*) This is method (*b*) under 'Learning,' again employed to prevent overweighting by extreme scores.

Relearning.—Shocks were not administered during this part of the experiment. 'Shocked errors' therefore refer to errors shocked during learning.

(*a*) See method (*a*) under 'Learning.'
(*b*) See method (*b*) under 'Learning.'
(*c*) See method (*c*) under 'Learning.'

(*g*) Here we calculated savings scores for shocked and non-shocked errors separately for each individual, and determined the critical ratio of the difference between the mean shock savings and the mean non-shock savings. Savings scores were calculated by means of the formula

$$\text{Savings} = 1 - \frac{N_r}{N_l},$$

in which

N_r = the number of (shocked or non-shocked) errors to relearning, and
N_l = the number of (shocked or non-shocked) errors to learning.

A summary of directions of differences and critical ratios of differences is presented in Table II.

Discussion

Considering first the analysis of the motor maze data presented in Table II, it will be noticed that only one of the 12 computed differences has a critical ratio over two. This difference, which indicates that more individuals have a higher rate of elimination of shocked than non-shocked errors, has a critical ratio of 2.43. As usually interpreted, this means that there are 99.2 chances in 100 that if the experiment were repeated a difference would be found

TABLE II

SUMMARY OF RESULTS OF ANALYSES

Condition	Method of Analysis	Kind of Errors Eliminated Most Rapidly	Critical Ratio	Chances in 100*
(Verbal Maze)				
Learning	*a*	Non-shock	1.03	85
	b	Non-shock	2.70	98
	c	Non-shock	2.16	99.7
	d	Non-shock	0.04	52
	e	Non-shock	2.20	98.6
	f	Non-shock	1.77	96
Recall	*a*	Non-shock	2.64	99.6
	b	Non-shock	3.28	99.9+
Relearning	*a*	Non-shock	2.04	98
	b	Non-shock	2.07	98
	c	Non-shock	2.57	99.5
	g	Non-shock	1.00	84
(Motor Maze)				
Learning	*a*	Shock	0.07	53
	b	Non-shock	0.73	77
	c	Shock	0.50	69
	d	Shock	0.01	50
	e	Shock	2.43	99.2
	f	Shock	0.19	58
Recall	*a*	Non-shock	0.47	68
	b	Non-shock	1.84	96
Relearning	*a*	Shock	0.58	72
	b	Shock	1.95	98
	c	Shock	0.88	81
	g	Shock	0.76	77

* That a repetition of the experiment under the same conditions would reveal a difference greater than zero and in the same direction.

larger than zero and in the direction indicated. Despite this single fairly significant difference, examination of Table II forces us to the conclusion that in this experiment there is no significant difference between the number of shocked errors and the number of non-shocked errors in learning. Furthermore there is no significant difference in recall or relearning between the number of errors on blinds shocked during learning and the number of errors on blinds never shocked.

Considering next the verbal maze, we find perfect agreement by all methods of analysis as to the direction of the difference. In learning, the errors which are shocked are eliminated less rapidly than those which are not shocked. The difference is not great in

comparison to the SD's, as shown by the low critical ratios. In recall and relearning the difference is also in the direction of making more errors which were shocked during learning than errors which were never shocked. Eight of the 12 differences have a critical ratio above two, and one (method 'b' in Recall) has a critical ratio of 3.28. These facts appear to justify our conclusion that in the verbal maze there is more successful avoidance during learning, recall, and relearning trials of errors which had *not* been shocked during learning than of errors which had been shocked. This is the direct converse of the difference found in the course of our previous study, in which it was found that shock resulted in more successful avoidance of shocked than non-shocked errors during the learning of a stylus maze problem. Our explanation of the present finding is as follows.

In the verbal maze, the subject had to make a choice between the members of a pair of nonsense syllables and pronounce his choice aloud. If the choice was correct, the next pair was presented. If the choice was wrong, he was either given a shock or told 'Wrong,' whereupon he pronounced the other syllable. The electric shocks in this situation appeared to have considerably more of a disorganizing effect than the shocks in the motor maze, presumably for two reasons: (1) The physical tension aroused by the expectancy of shock had no adequate outlet. Whereas the subject stood at the motor maze, he sat at the verbal maze. Whereas he could react to shock in the motor maze by a vigorous withdrawal movement, and usually did, there was no opportunity for a strong relevant reaction in the verbal maze. (2) The shock was not so obviously an integral part of the situation in the verbal as in the motor maze. In the motor maze the subject *gave himself* the shock in connection with an incorrect move, and it was received immediately upon making contact with the end of a cul-de-sac. In the verbal maze, the experimenter quite obviously administered the shocks. The length of time between the pronunciation of an incorrect syllable by the subject and the word 'Wrong' or the shock given by the experimenter varied considerably despite efforts to hold it as constant as possible, and many subjects reported vividly the disorganizing effect of the tense, agonized instant after they had pronounced a syllable when they sat and waited for the shock that might or might not come.

The shock, when it came, was disorganizing both because of the fearful expectancy which preceded it and the lack of a constructive outlet for the tension it aroused. The tension had to be dissipated in some way, and the result after a shock was frequently wandering of attention, readjustment of grip on the electrodes, a nervous laugh, or an irrelevant remark. This constituted distraction, or 'disruption' as Tolman, Hall, and Bretnall have termed it (7), and might be

expected to result in inferior learning of the shocked as compared with the non-shocked syllable pairs in the verbal maze. This we find to be the case.

With the possibility in mind that the difference in the results of the present experiment and our previous study (1) might be due in part at least to the absence of women from our present group of subjects, we reanalyzed the data of our previous experiment in search of sex differences. No evidence was uncovered of greater specificity of the effect of shock in either sex as opposed to the other. Apparently the sex difference in the composition of the groups is not the factor responsible for the difference in results.

The only other difference between our previous and the present experiment to which the difference in results could be logically attributed is the difference in the problems used. We have already discussed our hypothesis as to the reason for the direction of the shock specificity in the verbal maze, that is, why shock for certain errors and not for others resulted in the less rapid elimination of shocked than non-shocked errors. Now let us compare the motor maze situations of the present and previous experiments to determine the possible reason or reasons for the absence of shock specificity in the present study. The pattern of the maze used in this experiment is given in Fig. 1. It will be observed that each unit is identical with each other unit, save for the location of the blind. In the Fox maze (pattern illustrated in Bernard & Gilbert, 1), on the other hand, there is no such similarity of units. Each choice resembles each other choice only in that there are three directions of movement possible (excluding retracing, which was prevented), but in each case the pattern of the approach to the choice point is different from the pattern of the approach to each other choice point. This resulted in the turns in the Fox maze being learned individually by association with the pattern of the approach. Since the turns were learned individually, there was enhanced opportunity for the effect of shock for error to be restricted to a large extent to individual turns. It was observed that the subjects of the present motor maze experiment were unanimous in learning the turns by groups (*i.e.*, "One up, three down, two up," etc.) rather than individually. This grouping would result to a large extent in the loss of the separate identity of a unit, especially since a unit was identified by its temporal position in a series. With the decrease in unit individuality there was a corresponding decrease in the possibility of associating shock with a particular unit, and consequently a decrease in shock specificity. Furthermore when a subject lost count of his position while traversing the unidirectional maze, his subsequent performance during that trial was based either on chance or on a misconception of his position.

This happened frequently, and made for attenuation of our critical ratios. Such a lapse of memory in the Fox maze lasted only momentarily since the subject quickly recognized his position from the pattern of the path he was traversing.

There is another difference which may be partly responsible for the observed lack of shock specificity in the present motor maze experiment as opposed to the previous study. Because of the simplicity of the present pattern and the fact that each subject had the opportunity to practice on a sample maze of the same design (but without blinds), withdrawal from shock blinds was usually prompt and direct. Ordinarily only one shock was received, and then the circuit was made back to the correct path. In the Fox maze the choices were considerably more complex. Furthermore the subjects were not given any idea as to the design of the maze before the learning trials began. The result was that they frequently received several shocks in the same cul-de-sac on the same trial, especially during the early trials of learning, before they succeeded in locating the true path. This occurred despite the fact that retracing along the true path was prevented. These multiple entries were not recorded as such, but shock was given on every contact with a shock blind. The multiple shocks might well be expected to result in better memory in succeeding trials for errors so shocked, and thus augment the difference between the number of shocked and non-shocked errors made.

Conclusions

In our first study of the specificity of the effect of shock in learning (1), we had found shocked errors to be eliminated more rapidly than non-shocked errors in the Fox stylus maze situation. In the present study there was no apparent specificity in the unidirectional motor maze, and specificity in the opposite direction in the verbal maze—*i.e.*, non-shocked errors were eliminated faster than shocked errors. This leads us to the following conclusions:

1. The specificity of the effect of shock for error in learning situations may be in the direction of either selective retention or selective elimination of the shocked errors.

2. Such specificity, when present, may be significant for recall and relearning some days after learning has been completed.

3. The presence or absence of shock specificity and its direction depends on a number of factors, of which the type of problem is only one.

Summary

We have presented an experimental study of the specificity of shock for error in learning, planned along lines similar to a previous

study by the writer and Gilbert (1), but extended to include a stylus maze of a different pattern and a verbal maze. Data were gathered for learning, recall, and relearning by testing 62 undergraduate students at Clark University, the method being to administer shock for some but not all errors. The motor maze data show no significant difference between the number of shocked errors and the number of non-shocked errors in learning. Furthermore there is no significant difference in recall and relearning between the number of errors on blinds shocked during learning and the number of errors on non-shock blinds. The verbal maze data show more successful avoidance during learning, recall, and relearning trials of errors which had not been shocked during learning. Since these results differ from those of our previous experiment (in which we found that shocked errors were eliminated more rapidly than non-shocked errors in a stylus maze), possible reasons for the discrepancy are discussed.

We conclude that the findings of Bernard & Gilbert cannot be generalized. Although, as they found, shocked errors may be eliminated more rapidly than non-shocked errors in some learning situations, there are other situations in which shocked errors may be eliminated *less* rapidly, and still others in which there may be no difference between the rate of elimination of shocked errors and the rate of elimination of non-shocked errors.

(Manuscript received December 5, 1941)

References

1. Bernard, J., & Gilbert, R. W. The specificity of the effect of shock in maze learning with human subjects. *J. exp. Psychol.*, 1941, **28**, 178–186.
2. Crafts, L. W., & Gilbert, R. W. The effect of punishment during learning upon retention. *J. exp. Psychol.*, 1934, **17**, 73–84.
3. Hull, C. L. The meaningfulness of 320 selected nonsense syllables. *Amer. J. Psychol.*, 1933, **45**, 730–734.
4. Luh, C. W. The conditions of retention. *Psychol. Monogr.*, 1922, **31**, No. 3.
5. Miles, W. R. Unidirectional mazes for studying human learning. *J. gen. Psychol.*, 1931, **5**, 269–275.
6. Peterson, J. The backward elimination of errors in mental maze learning. *J. exp. Psychol.*, 1920, **3**, 257–280.
7. Tolman, E. C., Hall, C. S., & Bretnall, E. P. A disproof of the law of effect and a substitution of the laws of emphasis, motivation and disruption. *J. exp. Psychol.*, 1932, **15**, 601–614.

THE EFFECT OF TEMPERATURE ON THE RETENTION OF A MAZE HABIT IN FISH[1]

BY JOHN W. FRENCH

Princeton University

This experiment was designed to determine the effect of temperature on the process of forgetting. Cold-blooded animals are best suited as subjects in this type of investigation, since the temperature of their bodies can be controlled over a wide range. Goldfish were used, because they have proved to be fairly able learners in discrimination and conditioning situations. Thorndike (**13**) and Churchill (**1**) found that goldfish could learn a simple type of maze. In the present experiment a 4-blind linear maze was used.

The effect of temperature upon many physiological reactions has been studied. Crozier (**2, 3**) has shown positive relations between temperature and the velocity of progression of snails, the rate of chirping of tree crickets, the frequency of flashing of fireflies, and the rate of breathing in goldfish. Gasser (**5**) has shown the velocity of the nerve impulse to have a temperature coefficient similar to that of a chemical reaction, while the coefficient for the amplitude of the nerve impulse resembles that of the viscosity of fats.

Reactions which are more purely psychological have also been studied. François (**4**) and Hoagland and Perkins (**7**) have shown that time perception, as measured by the rapidity of tapping at a required rate, has a coefficient like that of a chemical reaction. Kleitman, Titelbaum and Feiveson (**10**) in measuring reaction times have shown that a rise in body temperature has a more pronounced speeding effect upon complex than upon simple reactions. They concluded that the mental processes involved in the complex reaction time problems are either of a chemical nature or are affected by the rate of metabolism in the cells of the cortex.

Two experiments dealing expressly with the effect of temperature upon retention are those of Hunter (**8**) and Hoagland (**6**). Hunter studied the effect of temperature on the retention of a darkness-avoiding response in cockroaches. Relearning of the response was unfavorably affected by intense cold, but this result was attributed to injury to the animals rather than to an effect on retention. Hoagland found retention in ants to be unaffected by exposure to tempera-

[1] From the Psychological Laboratory of Princeton University.

tures between 15° C. and 25° C. Above that range injury seemed to occur and made the results on retention difficult to interpret.

Experimental Procedure

In advance of the main experiment the fish were given a preliminary training of 12 trials in a short dummy maze to insure adaptation to handling. Then in the actual experiment the maze shown in Fig. 1 was used. The fish were poured from a glass tumbler into the starting box at the left. The correct path to the goal involved four choice points, two right turns and two left turns. Glass doors indicated by the dotted lines were closed whenever the fish were about to pass through them in a reverse direction. Stones were strewn over the bottom, since preliminary tests showed such cues to facilitate learning. The water in the maze was about three inches deep; its temperature was always kept at 22° C. The alleys were one inch wide.

Several conditions were provided to motivate the fish. First, since Perkins (12) found medium intensities to be preferred by goldfish, the goal was kept shaded while the rest of the maze was strongly illuminated. The illumination was supplied by a 100-watt lamp in a reflector 8 inches from the surface of the water. Secondly, a social motivation may have been provided

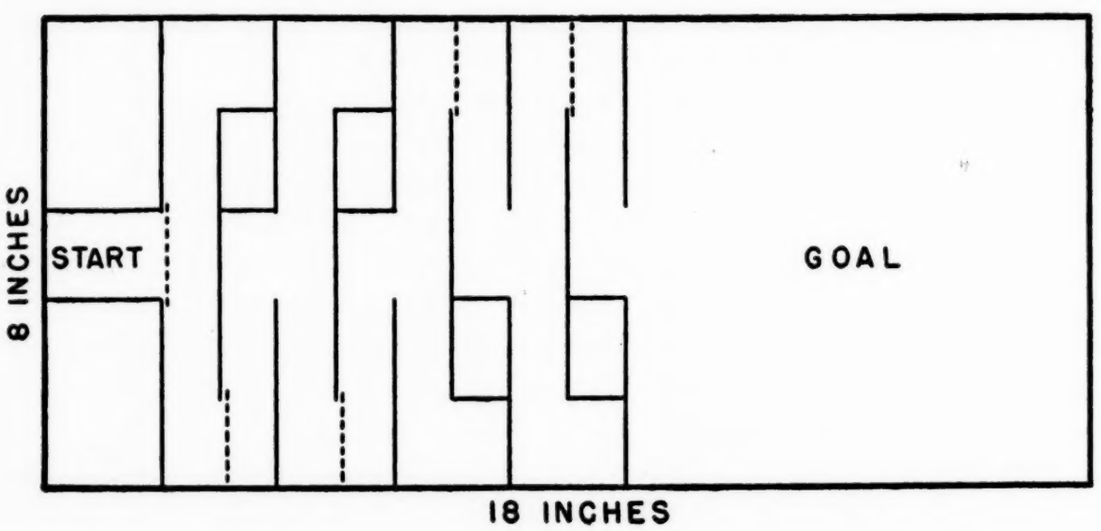

Fig. 1. Four-blind linear maze used for the original learning and relearning tests.

by the fact that two fish were kept in the goal box while a third was running the maze. Finally, since the fish were 1½ to 2 inches long, it was difficult for them to turn around whenever they strayed from the correct path. Hunger could not be used as a motivation, because the smallness of the fishes' appetite made it impossible to feed them after every trial. Nor was it advisable to withhold food entirely, since hungry fish were found in preliminary tests to continue exploration of the blinds long after learning was complete. For that reason they were always fed just before the learning period.

A fish was given one trial every four minutes. An error was scored whenever it entered a blind far enough to admit any part of the tail. The learning period was continued until a criterion of five consecutive errorless trials was reached. This occurred on the average after 30 trials, during which an average of 43 errors was made. The fish was then placed in a quart bottle containing water at 22° C., the temperature of the maze. The bottle was placed in a container which was maintained at one of the three temperatures, 28°, 16°, or 4° C., until the next day. The temperature of the water was allowed to reach these temperatures gradually so as not to harm the fish. Complete darkness was maintained during the period of controlled temperature.

After twenty hours at a controlled temperature the fish was brought back to 22° C. over a 15-minute period by slowly adding water at 22°. It was then put into the goal compartment of

the maze, where it was given $1\frac{3}{4}$ hours to adapt to the new temperature. As on the first day, the fish was fed before the day's experimentation began.

Following the same procedure as on the previous day, relearning was carried out until the original criterion of five consecutive errorless trials was reached. The number of errors made during the relearning period is taken as a measure of the forgetting which occurred since the original learning. The same procedure was followed for three more days, so that four relearning scores were obtained for each fish.

The temperature groups were equated for learning ability in terms of the error scores during the original learning. Although the range for the individual fish was from 12 to 100 errors, the equating process made the averages for the groups nearly alike, namely, 42.1, 42.9, and 42.3 for the 28°, 16°, and 4° groups, respectively. In all, 24 fish were included in the 28° group, 12 in the 16° group, and 12 in the 4° group.

Results

The results are shown in Fig. 2. The three curves represent the three temperature groups. The abscissa shows the five days of ex-

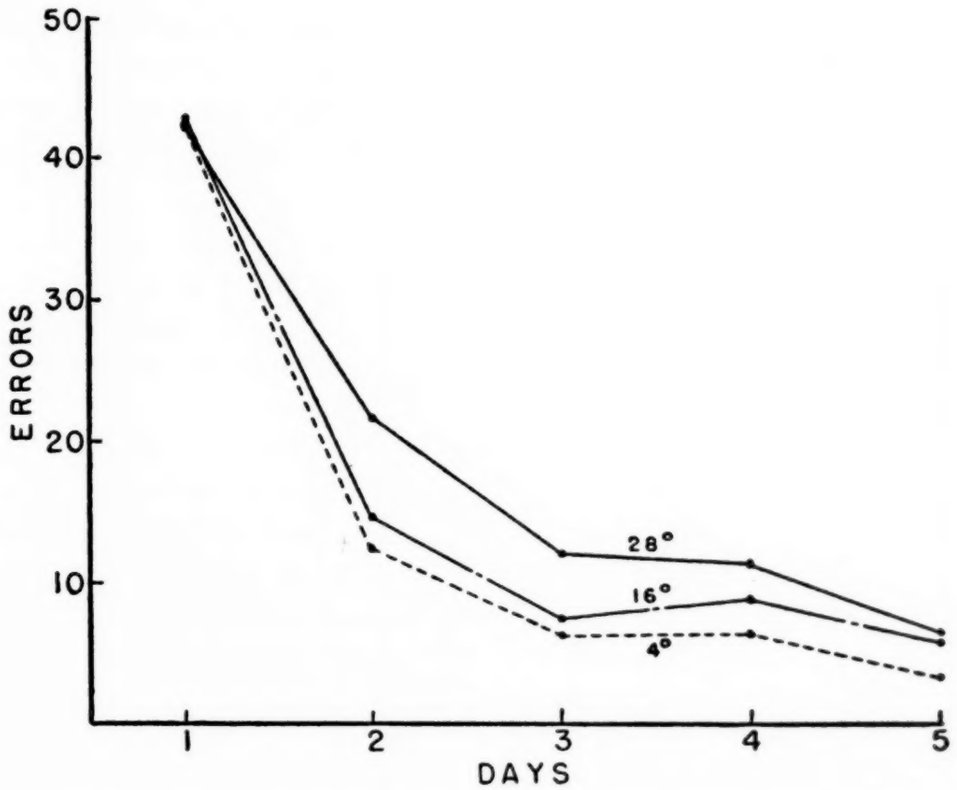

Fig. 2. Average number of errors made by each of the three temperature groups during the original learning (day 1), and the four relearning tests.

perimentation, and the ordinate shows the average number of errors made on each day. The results for the first day represent the average score made in the original learning. The last four days show the relearning scores. On every day the group that had spent the retention period at 4° made the fewest errors in relearning, while the group that had spent the retention period at 28° made the most. The average error scores and their standard deviations are shown in Table I.

TABLE I

AVERAGE ERRORS MADE BY THE THREE TEMPERATURE GROUPS DURING THE ORIGINAL LEARNING AND THE FOUR DAYS OF RELEARNING

Groups	Number of Fish	Day 1		Day 2		Day 3		Day 4		Day 5	
		M.	S.D.	M.	S.D.	M.	S.D.	M.	S.D.	M.	S.D.
4	12	42.3	22.9	12.3	7.2	6.4	4.0	6.3	6.3	3.3	3.4
16	12	42.9	21.3	14.6	10.8	7.4	6.1	8.9	6.7	5.8	3.4
28	24	42.1	16.9	21.6	11.4	12.0	5.8	11.3	6.6	6.4	4.1

A general score representing the relative amount of forgetting during an interval spent at each temperature was desired. The errors made by each fish on the four days of relearning were totalled, so that a total error score for each fish was obtained. A group average of these total error scores was then calculated. The fish kept at 28° during the retention interval made an average total error score of 51.3, those kept at 16° a score of 36.8, and those kept at 4° a score of 28.3. These figures are represented by the solid line curves in Figs. 3 and 4.

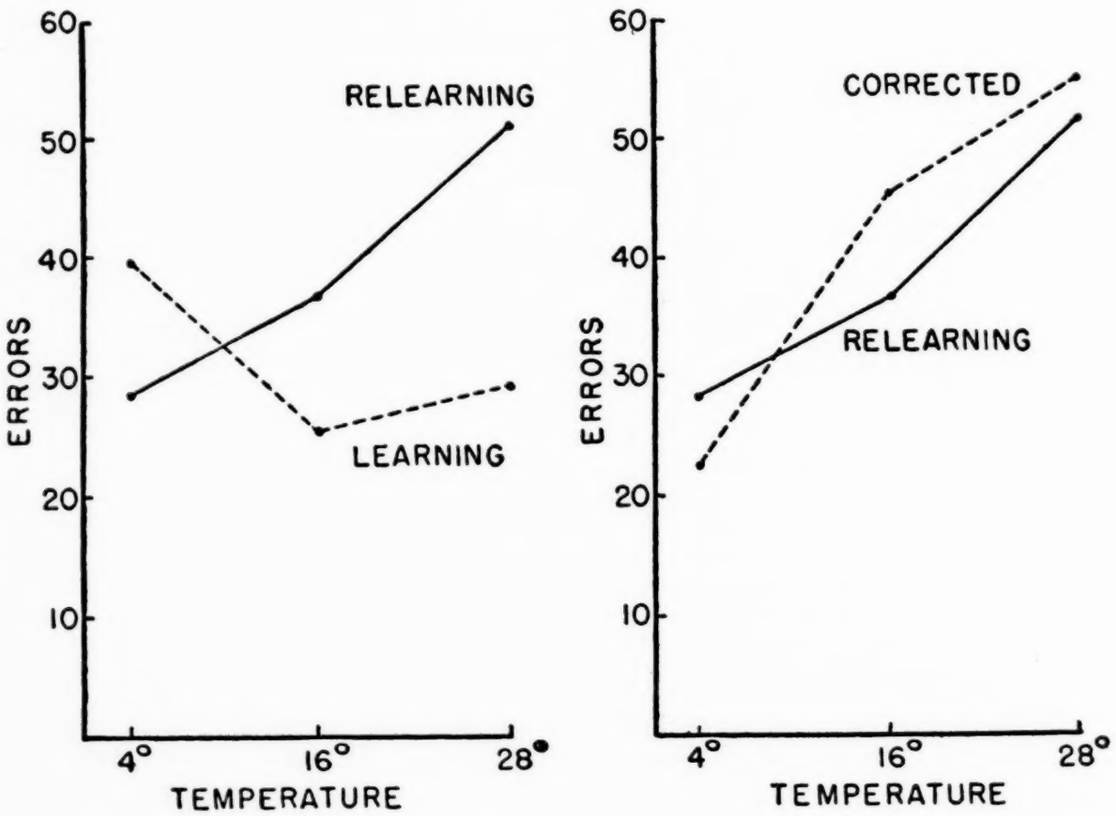

FIG. 3. Comparison of the effect of different temperatures upon maze performance with its effect upon relearning.

FIG. 4. Comparison of relearning data with the same data after correction for the direct effect of exposure to temperature.

Fisher's *t*-method was used to ascertain the significance of the differences obtained. For the extreme groups $t = 5.02$, which indicates a significance well below the 1 percent level. The difference between the 16° and 28° groups is just significant at the 1 percent level; $t = 2.67$. The difference between the 16° and 4° groups is far less significant; $t = 1.39$.

Trials were treated in the same way as errors. The number of trials required for relearning on the four days was totalled for each fish. An average total score was then computed for each of the three groups. The differences are all in the same direction as those between the average total error scores. Fisher's *t* for the difference between the extreme groups is 5.35. Time scores were found to be unreliable. They ranged from 1 to 8 minutes at the beginning of the learning and from 10 seconds to 2 minutes at the end.

Discussion

The results suggest that retention varies inversely with water temperature. However, the question arises whether these results are in fact a function of retention. It is conceivable that twenty hours in water near freezing or near the maximum bearable temperature may have some direct effect upon the learning ability of the fish. The effect of temperature on learning might result from a variation in metabolism, mobility, or hunger. Any such variation could account for the differences in the relearning scores.

To test this possibility, it was necessary to measure original learning after a period of exposure to different temperatures. Groups of 12 fish were put in water at 4°, 16°, and 28° for twenty hours under the same conditions as in the main experiment. The fish were then brought to 22° over a 15-minute period, and given $1\frac{3}{4}$ hours to adapt. They were given twelve trials in the dummy maze, and then were taught the principal maze until the criterion of five errorless trials was reached.

The average number of errors made in learning the maze after 20 hours at each of the three temperatures is shown in the dotted curve of Fig. 3. The cold group was the slowest to learn, averaging 39.3 errors. The warm group followed, making 29.2, and the medium group learned most easily, making an average of only 25.3 errors.[2] These data were treated by a simple analysis of variance as described by Lindquist (11). The ratio between the variance for temperature groups and the variance for within groups is 5.27. For the number

[2] These averages are all less than the average number of errors made on the first day of the main experiment. This difference is probably due to the stock of fish used. They were obtained from the same supply house, but at a later time in the year.

of degrees of freedom involved, this indicates a significance between the 1 percent and 2 percent levels.

Since exposure to cold did not affect learning in the same way that it affected relearning, it is concluded that the relearning results are in fact a measure of retention. The period of exposure to a controlled temperature does affect the maze score, but its effect is such as to increase rather than decrease the significance of the differences shown in relearning. The cold group thus may be said to show the best retention in spite of the detrimental effect of intense cold on maze performance.

In order to arrive at the true relation between temperature and retention, it is necessary, in view of the facts just presented, to compensate for the direct effect of temperature upon maze performance. The mean error score of the three groups in the check experiment was taken as the normal number of errors necessary to learn. The deviations from the mean were considered the result of exposure to the controlled temperature. A correction factor for each group to compensate for its deviation was arrived at by dividing the mean score by the group score. The original scores of the main experiment were then corrected by multiplying by the corresponding correction factor. The solid curve in Fig. 4 again represents the original relearning data. The dotted curve represents the same data after compensation for the direct effect of temperature on factors other than retention.

In view of the experiments of Jenkins and Dallenbach (**9**) and van Ormer (**14**), which showed that retention is better after a period of sleep or inactivity than it is after waking and activity, a second check experiment became necessary. It was conceivable that the results were due to greater activity at the warmer temperatures, which might alter retention through retroactive inhibition.

In order to test this possibility, an apparatus for measuring the activity of fish was constructed. A series of celluloid fins was placed in the bottle with the fish. These fins were so arranged that movements of the animal would make and break a contact which activated a counter. In this way the relative activity of the fish during the following three periods was tested: (1) the period of controlled temperature in the dark; (2) the $1\frac{3}{4}$-hour adaptation period; and (3) the period in which the fish would ordinarily have been occupied in relearning. The activity records of eight fish were obtained for two successive days under each of the three temperature conditions.

The amount of activity during period (1) is shown in Fig. 5. The solid line curve represents the corrected retention scores, as shown on the left hand ordinate. The dotted curve shows the activity in number of counts, as shown on the right hand ordinate. The fish

were most active at 16°, making an average of 12,764 counts. At 28° they made 4,996, and at 4° they made 1,608.[3] Since the activity of the same eight fish was tested under the three temperature conditions, the *t*-test for related means is applicable to these data. For the difference between the 4° and 28° groups $t = 2.43$; for that between the 4° and 16° groups $t = 4.45$; and for that between the 16° and 28° groups $t = 3.64$. The excess of activity of the 16° group over the other two groups is significant at the 1 percent level.

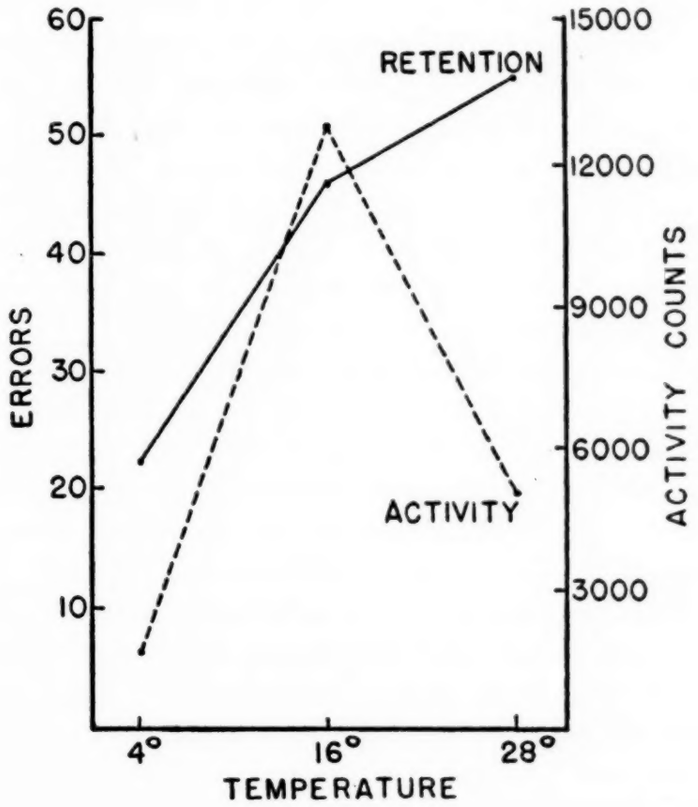

FIG. 5. Comparison of retention with amount of gross activity during the period of forgetting.

If retroactive inhibition due to activity plays an important part in retention, forgetting should be greatest after a rest period spent at 16°, as the fish were most active at this temperature in both periods (1) and (2). Since the results on retention do not show this trend, it must be concluded that retroactive inhibition due to general activity is not responsible for the results.

Just after feeding, the activity of the 16° group dropped below

[3] These results were taken under the specific conditions of the experiment. It is possible that the 16° group may have been subject to slight vibration, since the temperature for this group was maintained by keeping the experimental bottles immersed in a tank through which water was slowly running. The 28° and 4° temperatures were provided by an incubator and electric refrigerator. Since this was also true of the main experiment, these results are a true measure of the activity which may have affected the relearning scores.

that of the 28° group, so that in period (3) the activities of the three groups bore to each other the same rank order relation as did the relearning scores; the counts were 3,357, 1,574 and 358 for the 28°, 16°, and 4° groups, respectively. If these differences in activity had any effect on maze performance, it would have shown up in the results of the first control experiment. Perhaps it did, inasmuch as the poor learning performance of the cold group may have been in part caused by the animals' lethargic condition, as shown by the low count of 358.

In most of the experiments cited in the introduction a temperature coefficient was calculated. However, such a calculation is impossible in the present experiment for two reasons. First, it is doubtful that a temperature coefficient for retention would have much meaning, since maze errors are not units of equal magnitude. For example, a fish making twice as many errors as another in relearning a maze probably has not forgotten twice as much. Secondly, each temperature group was subject to a certain constant amount of forgetting outside the controlled temperature period. In the important period immediately after learning all the fish were disturbed by being poured into a bottle and carried to the ice-box or incubator. Then they only gradually reached the controlled temperature. Before relearning there was an adaptation period of $1\frac{3}{4}$ hours when all three groups were at 22° C. Since the extent to which processes during these intervals affect the relative differences between the retention scores is unknown, a temperature coefficient cannot be calculated.

It is often maintained that forgetting is produced by retroactive inhibition due to bodily or mental activity. It has been shown in this experiment that gross activity has little effect upon retention. Furthermore, we can assume that the mental life of a lower animal like the fish is decidedly limited when the animal is isolated in a smooth-walled container in the dark. Forgetting, therefore, cannot be simply a function of inhibition. In fish, at least, the essential process is a gradual fading of the impressions established in the course of learning. This experiment has shown that the rate of such fading varies with temperature.

Summary and Conclusion

In order to test the effect of temperature upon retention, goldfish were taught a maze at 22° C. Some of these were then maintained for a rest period of about a day at each of three temperatures, 28°, 16°, and 4° C. Retention was then measured by relearning tests at 22° C. The group which had spent the retention period at 28° made almost twice as many errors during relearning as the group which

had spent the retention period at 4°. Relearning scores for the group kept at 16° show that the greater part of the difference comes between 4° and 16°, indicating a negatively accelerated temperature function.

Two control experiments were carried out. One of these showed that the results were due to the action of temperature upon retention and not to its action upon the rate of learning. The other showed that the differences were not due to retroactive inhibition produced by general activity during the period of controlled temperature.

It is concluded that there exists a forgetting process which is independent of inhibition and whose rate is a function of temperature.

(Manuscript received November 25, 1941)

References

1. Churchill, E. P. The learning of a maze by goldfish. *J. anim. Behav.*, 1916, **6**, 247–255.
2. Crozier, W. J. On the critical thermal increment for the locomotion of a diplopod. *J. gen. Physiol.*, 1924–1925, **7**, 123–136.
3. Crozier, W. J., & Stier, T. J. B. Critical increment for opercular breathing rhythm of the goldfish. *J. gen. Physiol.*, 1924–1925, **7**, 699.
4. François, M. Contribution a l'étude du sens temps. La température interne comme facteur de variation de l'appréciation subjective des durées. *Année psychol.*, 1927, **28**, 186–204.
5. Gasser, H. S. Nerve activity as modified by temperature changes. *Amer. J. Physiol.*, 1931, **97**, 254–270.
6. Hoagland, H. A study of the physiology of learning in ants. *J. gen. Psychol.*, 1931, **5**, 21–41.
7. Hoagland, H., & Perkins, C. T. Some temperature characteristics in man. *J. gen. Physiol.*, 1935, **18**, 399–408.
8. Hunter, W. S. The effect of inactivity produced by cold upon learning and retention in the cockroaches. *J. genet. Psychol.*, 1932, **41**, 253–266.
9. Jenkins, J. G., & Dallenbach, K. M. Obliviscence during sleep and waking. *Amer. J. Psychol.*, 1924, **35**, 605–612.
10. Kleitman, N., Titelbaum, S., & Feiveson, P. The effect of body temperature on reaction time. *Amer. J. Physiol.*, 1938, **121**, 495–501.
11. Lindquist, E. F. *Statistical analysis in educational research.* Cambridge: Riverside Press, 1940, 93.
12. Perkins, F. T. A further study of configurational learning in goldfish. *J. exp. Psychol.*, 1931, **14**, 508–538.
13. Thorndike, E. A note on the psychology of fishes. *Amer. Nat.*, **33**, 923–925.
14. Van Ormer, E. B. Retention after intervals of sleep and waking. *Arch. Psychol.*, 1932, **21**, No. 137, 1–49.